

# Jerry Vishal Joseph Arockiasamy

📍 Fairfax, VA 📞 +1-571-774-0022 ✉️ [jjosepha@gmu.edu](mailto:jjosepha@gmu.edu) [in](#) [LinkedIn](#)

## Skills

- **Technical & Data Analysis:** SQL, Python, Fraud Detection, Anomaly Detection, Data Quality Monitoring, Statistical Methods
- **Business Intelligence & Tools:** Tableau, PowerBI, Figma (Dashboard Design), ArcGIS (Spatial Analytics), Business Intelligence Tools
- **Big Data & Cloud:** Databricks, AWS, Azure
- **Machine Learning & Feature Engineering:** Scikit-learn, Pandas, NumPy, Trend Identification, Summary Statistics, ML Modelling

## Experience

### Amazon | *Data Analyst*

2022 - 2023

#### Data Analyst – Reporting & Automation (Jan 2023 – Jun 2023)

- Engineered automated SQL and VBA data pipelines, reducing manual workload by 30% and enabling real-time monitoring of (KPIs).
- Managed Tableau dashboards to visualize business metrics, enhancing decision-making and data visibility for cross-functional teams.
- Ensured data accuracy by validating new and existing tools, performing data audits, and implementing cleansing processes, which improved data integrity and streamlined reporting workflows.

#### Data Analyst – Fraud Detection (Oct 2022 – Jan 2023)

- Analyzed large datasets to identify fraud patterns and anomalies, improving reporting accuracy by 20% and reducing fraudulent transactions using Python-based validation checks and SQL queries.
- Optimized machine learning-driven fraud detection models, decreasing fraudulent transactions by 15% and enhancing risk mitigation strategies within e-commerce pipelines.
- Collaborated with cross-functional teams (engineering, business intelligence) to integrate fraud detection models into production systems, ensuring alignment with business objectives and enhancing user safety.

### Mentored Minds | *Data Analytics Intern*

2022

- Led end-to-end **ETL pipeline** development, defining project scope, deliverables, and timelines. Coordinated with cross-functional teams to align workflows with business goals, reducing project completion time by 20%.
- Automated monthly sales report generation by developing a Python script, saving 8 hours per week, and enabling the sales team to focus on upselling opportunities, not reporting.
- Led data quality initiatives that uncovered and resolved three major data integrity issues, eliminating reporting discrepancies across 10,000+ records and strengthening executive decision-making.

### Tata Group | *Data Analytics Intern*

2021 - 2022

- Transformed 5+ datasets (sales, inventory, demand forecasting), improving data quality by 40% and reducing processing time by 15%.
- Incorporated Power BI dashboards with custom DAX measures to track KPIs like order fulfillment and inventory turnover.
- Enabled stakeholders to identify bottlenecks, boosting operational efficiency by 25% and enhancing safety for high-risk logistics.

## Education

### George Mason University

2023 - 2025

*Master of Science, Data Analytics Engineering (GPA: 3.82)*

- **Coursework:** Machine Learning, Big Data Analytics, Statistical Methods, Data Mining

### Loyola ICAM College of Engineering and Technology

2017 - 2021

*Bachelor of Engineering, Electronics and Communication*

## Projects

### Real-Time E-commerce Transaction Data Streaming and Processing Using Modern ETL Tools

2025

- Developed an end-to-end real-time data pipeline using Apache Airflow, PostgreSQL, and AWS Lambda to extract, transform, and load e-commerce transaction data for real-time analytics and fraud detection.
- Incorporated Docker and Kubernetes for scalable deployment of the pipeline, achieving a 30% improvement in processing efficiency and enabling seamless handling of high-volume transaction data.
- Integrated data from disparate sources to ensure seamless ETL workflows and optimized query performance for real-time insights.

### Amping Up Insights: Exploring Electric Utility Rate Trends

2024

- Leveraged K-Means clustering to categorize 500+ utility providers, pinpointing five vendors offering rates 15% below market average across industrial, commercial, and residential sectors, resulting in \$20,000 cost savings.
- Employed SQL and Tableau to detect **billing anomalies** across 10,000+ utility records, identifying cost-saving opportunities for clients.

### Sentiment Analysis in Customer Reviews

2024

- Extracted and processed customer reviews using AWS Glue DataBrew, identifying recurring product flaws and flagging negative feedback to improve issue resolution time by 30%. Analyzed key trends and anomalies to enhance data-driven decision-making and refine product strategies.
- Constructed features using TF-IDF scores and VADER sentiment analysis; produced summary statistics like review length distribution to pinpoint key trends in customer feedback, increased model accuracy by 25%.

## Certifications

- **Professional Certificate on Google Data Analytics:** 2022
- **Hands-on Project: Data Visualization for Business Insights:** 2022
- **Learning Excel: Data Analysis:** 2022