# Jerry Vishal Joseph Arockiasamy

Fairfax, VA

+1 (571) 774-0022 - jjosepha@gmu.edu - linkedin.com/in/jerryvishal

## Professional Summary

Results-oriented Data Analyst with hands-on experience at Amazon, where I drove a 15% reduction in operational inefficiencies and improved data management by 20%, resulting in $50,000 annual savings. Skilled in SQL, Python, Tableau, ETL pipeline development, with a strong foundation in predictive analytics, data visualization, and statistical analysis. Currently pursuing a Master's in Data Analytics Engineering at George Mason University. Google Data Analytics Certified professional with a proven track record of using data-driven insights in the process of strategic decision-making.

## Skills

**Programming Languages:** Python - R - SQL - HTML - CSS – C, **Frameworks:** React - Django – WordPress, **Developer Tools**: Git - AWS - Google Cloud Platform - Databricks - VS Code - PyCharm - Eclipse - Power Query, **Data Visualization Tools:** Tableau - Power BI - QlikView - RapidMiner - SAS Data Visualizing Tool, **Data Management:** Microsoft SQL Server - Oracle - MongoDB - SAP Data Cloud, **Analytical Skills:** Predictive Modeling - Statistical Analysis - Data Cleaning - Dashboard Development, **Proficiencies:** Financial Data Analysis - Regulatory Compliance - Financial-Reporting - Data Lifecycle Management

## Education

**George Mason University**                                                                                        Fairfax, VA
Master of Science in Data Analytics Engineering (CGPA: 3.78)                              Aug. 2023 – Exp. May 2025
(Machine Learning, Big Data Analytics, Statistical Methods, Data Mining)

**Loyola ICAM College of Engineering and Technology**                                        Chennai, India
Bachelor of Engineering in Electronics and Communication.                                 Aug. 2017 – May 2021

## Experience

**Mentored Minds**                                                                                                        Remote
Data Analytics Intern                                                                                   Mar. 2023 – Jun. 2023

- Automated ETL processes using Python and SQL, improving efficiency by 25%.
- Designed interactive dashboards and reports, aiding clients in data-driven decision-making.
- Built and maintained complex data models to optimize analytics workflows, reducing analysis time by 20%.
- Validated and cleaned large datasets, ensuring data integrity and seamless integration for stakeholders.

**Tata Group**                                                                                                            Remote
Data Analytics Intern                                                                                     Oct. 2022-Jun. 2022

- Extracted raw data from Excel and performed comprehensive cleaning, transformation, and normalization to ensure data integrity for analysis.
- Imported pre-processed data into Power BI, leveraging DAX functions and Power Query for advanced data modeling and calculations.
- Designed and developed dynamic, interactive dashboards and reports tailored to client specifications, enabling real-time insights.
- Presented key findings and actionable insights through data storytelling, highlighting critical trends and patterns to stakeholders.

**Amazon**                                                                                                            Bangalore, India
Data Analyst                                                                                              Dec. 2021 – Jun. 2022

- Led onboarding, access, and data management processes for Premium Seller partners, enhancing operational efficiency by 30%.
- Designed and deployed SQL-based ETL workflows, improving data pipeline performance by 25%.
- Developed and maintained advanced metrics, automated reports, and interactive dashboards using SQL and Tableau.
- Created new data warehouse tables using SQL, Python, and R, reducing processing time by 40%.
- Streamlined reporting operations, achieving 50% faster turnaround times through improved data visualization.
- Demonstrated expertise in analytical problem-solving by designing robust data solutions that addressed complex business challenges in a results-driven environment.

## Projects

**Amping Up Insights: Exploring Electric Utility Rate Trends** | Python, Tableau, Power BI, Random Forest, K-Means Clustering          Dec 2024

- Conducted in-depth analysis of electric utility rates across ownership models and geographic regions using Python and machine learning techniques.
- Designed and implemented clustering models (K-Means) to identify cost-effective utility providers and rate trends across sectors.
- Developed predictive models (Random Forest) to determine revenue drivers, achieving actionable insights for underperforming utilities.
- Created dashboards in Tableau and Power BI to communicate findings on rate disparities, ownership influence, and socioeconomic impacts.
- Addressed energy equity by identifying patterns in rate structures, emphasizing affordability and sustainability for underserved regions.

**Sentimental Analysis in Customer Reviews** | AWS, Databricks, PySpark, Python          Apr 2024

- Oversaw an in-depth exploratory data analysis (EDA) in AWS using AWS Glue Data brew; identified key data trends and anomalies, which led to a 25% improvement in data quality and accuracy.
- Performed sentiment analysis and topic modeling on a Data Frame containing processed review text and utilized the VADER lexicon for sentiment analysis and the LDA (Latent Dirichlet Allocation) algorithm for topic modeling.
- Used PySpark for analysis in Databricks and employed NLTK packages for extensive sentimental analysis.

**Experimental Approach to Detect Autism** | Power BI, Node-red, C, Python          June 2020

- Designed an IoT-enabled toy for real-time autism assessment, achieving 50% higher accuracy compared to traditional methods.
- Engineered an AI diagnostic model using machine learning to predict behavioral patterns with a 40% accuracy improvement.
- Integrated multi-sensor data streams into a centralized data pipeline for seamless analysis using Python and Node-RED.
- Conducted statistical analysis with R to ensure model compliance with clinical benchmarks.

## Certifications

- Professional Certificate on Google Data Analytics (Coursera, 2022)
- Hands-on Project: Data Visualization for Business Insights (Forage/Tata, 2022)
- Learning Excel: Data Analysis (LinkedIn, 2022)