Ritesh Somashekar

✉ ritesh.chandra33@gmail.com ↗   📞 +1 571-245-7770 ↗   📍 Fairfax, VA   **in**LinkedIn ↗   ⭘ Github ↗

Technologies

- **Programming Languages:** Python, SQL, R, JAVA, C
- **Data Engineering Tools:** PowerCenter, Precisely Data360, Precisely Assure, AWS (S3, Lambda, Athena), Databricks, PySpark, Microsoft SQL Server, MySQL
- **Data Analysis & Visualization:** Pandas, NumPy, Scikit-Learn, Tableau, Power BI, Excel (Formulas, VBA)
- **Other Tools:** Git, Jupyter, Microsoft Power Platform (Power Apps, Power Automate)
- **Techniques:** Data Pre-processing, ETL, Data Warehousing, Statistical Analysis, Data Visualization, Machine Learning, Web Scraping

Experience

Graduate Research Assistant ⸺ *Fairfax, VA*
Costello College of Business ⸺ *August 2024 – Present*

- Designed **Python and R-based algorithms** to **investigate social cohesion and community coping** behaviors **during disaster scenarios** using large-scale real-world datasets.
- Leveraged graph-based analytics to uncover pattern insights and utilized ensemble models to forecast trends, **achieving an RMSE of 0.89.**
- Visualized findings using **R and Tableau** to support stakeholder understanding and strategic research goals.

Data Engineer ⸺ *Bangalore,India*
Carelon Global Solutions (Elevance Health) ⸺ *May 2022 – Aug 2023*

- Automated **20+ agile** end-to-end ETL pipelines for **provider and consumer analytics**, focusing on **invoice balancing and dashboard reporting** using Python, Informatica PowerCenter, Precisely (Assure, Data360), and AWS.
- Built and deployed **web crawlers to parse over 400K XML/JSON-based healthcare reports**, enhancing data availability for downstream applications **with maximum efficiency and cutting manual overload by 90%.**
- Engineered **Python-based data quality scripts and complex SQL queries, improving efficiency by 60% in the Edward Profiling** project and reducing manual interventions. Developed **Tableau and Power BI dashboards** to present provider insights and performed root cause analysis for performance anomalies.

Associate ETL Engineer ⸺ *October 2020 – Apr 2022*

- Played a key role in the Seven Plus Locations project by **automating Exploratory Data Analysis & Data Profiling pipelines** using **Python** and visualizing healthcare record data in **Tableau and Power BI**.
- Developed algorithms for **automating medical invoice processing**, improving claims balancing accuracy and **reducing error rates by 70%.**

Education

George Mason University ⸺ *Aug 2023 – May 2025*
MS in Data Analytics and Engineering, GPA:3.93/4.0

Related Courses: Analytics and Decision Analysis, Business Analytics,,Data Mining, Viz using Tableau, Advance Machine Learning, Neural Networks , Natural Language Processing, Applied Statistics

Projects

Agentic AI-Based Ticket Automation System | Agentic AI, SecureGPT, AWS , JIRA, NIST Compliance

- Designed an autonomous AI system that monitors cybersecurity anomalies and auto-generates JIRA service tickets aligned with SLA policies.
- Implemented agent-to-agent handshakes to streamline ticket triage with 93% automation accuracy, ensuring production-readiness..

Analyzing Diabetes Dynamics using Machine Learning |Python, R, EDA, Web Scraping, MongoDB, Databricks

- Conducted cluster-based segmentation and applied time-series forecasting to identify pre-diabetic patterns in U.S. population groups.
- Achieved strong model performance with an RMSE of 0.87 and MAE of 1.45, highlighting predictive accuracy for healthcare intervention planning.