

# Rahul S

Thoothukudi, Tamil Nadu | +91 9787496120 | [subbiahrahul007@gmail.com](mailto:subbiahrahul007@gmail.com)

[linkedin.com/in/rahuls6120](https://linkedin.com/in/rahuls6120) | [github.com/rahuls6120](https://github.com/rahuls6120)

## PROFESSIONAL SUMMARY

Data professional with expertise in **Data Engineering, ETL Development, and Analytics**, currently supporting production-grade healthcare data pipelines at Cognizant. Experienced in monitoring 55+ daily ETL jobs processing millions to billions of records using IBM DataStage, Informatica, and Oracle databases. Proficient in Python, SQL, Power BI, and cloud technologies (AWS), with hands-on experience in building ML models, developing interactive dashboards, and optimizing data workflows. Published researcher in AI/ML with strong foundation in transforming complex data into actionable business insights.

## PROFESSIONAL EXPERIENCE

### • Cognizant Technology Solutions

*Analyst Trainee - Data Engineering* Jun 2025 - Present  
Chennai, Tamil Nadu

- Supporting production ETL pipelines in healthcare domain, monitoring 55+ daily jobs processing millions to billions of records using IBM DataStage and Informatica, ensuring timely data delivery for business-critical operations.
- Performing source-to-target data validation and quality checks using Oracle SQL queries to maintain data integrity across enterprise systems.
- Managing batch job workflows via Autosys and monitoring server operations through PuTTY on Linux environments, contributing to operational reliability.
- Collaborating with cross-functional teams to analyze job failures, identify root causes, and implement efficiency improvements for time-sensitive data loads.
- Completed intensive 3-month training program covering IBM DataStage, Informatica, SQL, Python, AWS fundamentals, Linux administration, ETL design patterns, and Data Warehousing concepts.

### • GEP Solutions

*Data Analytics Intern* Dec 2024 - Mar 2025  
Coimbatore, Tamil Nadu

- Completed structured campus-to-corporate training program focusing on professional communication, team collaboration, and analytics tools.
- Analyzed procurement datasets using Excel and Power BI, generating insights to support data-driven business decisions.
- Contributed to data cataloging initiative, improving data organization, discoverability, and accessibility across the organization.
- Applied technical skills through hands-on assignments involving data cleaning, transformation, and visualization techniques.

## TECHNICAL SKILLS

- **ETL & Data Engineering:** IBM DataStage, Informatica (Monitor, Workflow Designer), ETL Pipeline Design, Data Warehousing, Autosys, Linux/Unix
- **Databases & Query Languages:** Oracle, MySQL, SQL (Advanced Queries, CTEs, Stored Procedures), Data Validation
- **Programming & Scripting:** Python (Pandas, NumPy, Scikit-learn, TensorFlow), SQL, Data Manipulation
- **Data Visualization & BI:** Power BI, Tableau, Excel (Advanced), Matplotlib, Seaborn, Dashboard Development, KPI Reporting
- **Cloud & Platforms:** AWS (EC2, S3, IAM - Fundamentals), Google Cloud, Google Colab, Jupyter Notebook
- **Machine Learning & AI:** Scikit-learn, TensorFlow, Hugging Face, Stable Diffusion, NLP, Prompt Engineering, Generative AI
- **Professional Skills:** Problem-Solving, Data Quality Analysis, Cross-functional Collaboration, Adaptability, Fast Learner

## TECHNICAL PROJECTS

---

### • Prompt Optimization for Text-to-Image Generation

Jan 2025 - Mar 2025



Python, Hugging Face, CLIP, Stable Diffusion, NLP, Google Colab

- Engineered an AI-powered system leveraging NLP techniques to automatically refine and optimize user prompts for text-to-image generation, improving output precision and relevance.
- Implemented dynamic model selection using Stable Diffusion based on prompt context analysis, enhancing generation quality and user experience.
- Published findings in peer-reviewed journal, contributing to research on prompt engineering methodologies.

### • Titanic Survival Prediction - Machine Learning

May 2024



Python, Scikit-learn, Pandas, Matplotlib, Jupyter Notebook

- Developed end-to-end ML pipeline for passenger survival prediction, achieving 80.36% accuracy through feature engineering, data preprocessing, and logistic regression modeling.
- Performed exploratory data analysis and visualization to identify key survival factors including passenger class, age, and gender.

### • Impact of ChatGPT on Student Learning - Data Analysis

May 2023



Python, Pandas, Excel, Data Visualization

- Conducted comprehensive analysis of survey responses to assess ChatGPT's influence on personalized learning, time management, and academic outcomes.
- Identified key trends, benefits, and limitations of AI-assisted learning through statistical analysis and data visualization techniques.

## EDUCATION

---

### • Sri Ramakrishna College of Arts & Science

Aug 2022 - May 2025

Coimbatore, Tamil Nadu

Bachelor of Science in Computer Science with Data Analytics

- CGPA: 7.87/10

- Relevant Coursework: Data Structures, Database Management, Machine Learning, Statistics, Python Programming

### • St. Mary's Boys Hr. Sec. School

Jun 2020 - May 2022

Thoothukudi, Tamil Nadu

Higher Secondary Certificate - Computer Science

- Percentage: 65.33%

## PUBLICATIONS

---

### [J.1] Prompt Optimization for Improved Text-to-Image and Image-to-Prompt Generation.

*International Journal of Research Publication and Reviews*, March 2025.

Mentor: Dr. Pranesh.

## CERTIFICATIONS & PROFESSIONAL DEVELOPMENT

---

### • IBMGenerative AI: Prompt Engineering Basics

Mar 2025

### • GoogleBuild Real World AI Applications with Gemini and Imagen

Jun 2025

### • AWS EducateIntroduction to Generative AI

Apr 2025

### • AWS Educate AWS Well-Architected Foundations

Feb 2025

### • Deloitte (Forage)Data Analytics Job Simulation - Tableau & Excel

Jun 2025

### • Google Cloud Skills BoostIntroduction to Generative AI

Nov 2024

### • TableauTableau Fundamentals

Sep 2024

### • AlteryxMachine Learning Fundamentals Micro-Credential

Aug 2024

**Additional Professional Learning:** Completed courses in Python Programming, SQL for Data Analysis, Cloud Computing Basics, Power BI Fundamentals, Cross-Cultural Communication, and Flexible Thinking Skills through Udemy and LinkedIn Learning (2024-2025).

## ACHIEVEMENTS & RECOGNITION

---

### • Best Participants - DIGITAL Women Conference & Mobility Programme

Feb 2024

Coimbatore, Tamil Nadu

Team Analyst

- Selected as best participants in competitive conference focusing on digital solutions for social impact.
- Led team of 6 members in analyzing survey data from 330+ respondents on women's safety and digital empowerment.
- Designed user interface prototype incorporating survey insights, emphasizing safety features and accessibility.

## LEADERSHIP & VOLUNTEER EXPERIENCE

---

### • Director - Editorial Board

*Rotaract Club of Sri Ramakrishna College of Arts & Science*

*Jun 2023 - Apr 2024*

*Coimbatore*

- Led visual content strategy and design for 15+ club events, managing poster creation and promotional materials.
- Documented club activities through photography and videography, building comprehensive event archive.
- Collaborated with cross-functional teams to ensure consistent branding and effective communication.

## ADDITIONAL INFORMATION

---

**Languages:** English (Fluent), Tamil (Native), Telugu (Conversational)

**Interests:** Photography, Gardening, Traveling, Technology Trends, AI/ML Research