

ARUNKUMAR S

📍 India

✉️ arunkumarse315@gmail.com | ☎️ +91 9750814876 | GitHub

PROFESSIONAL SUMMARY

Data Science aspirant with **3.7 years of experience in Software Testing (QA Automation)** and a strong transition focus into **Data Science & Analytics**. Hands-on experience in **Python, SQL, Machine Learning, Data Analysis, and Visualization** through real-world projects across music analytics, energy forecasting, mobility analytics, and digital payments. Strong foundation in **data validation, automation, problem-solving, and business-focused insights**, backed by industry exposure in large-scale enterprise systems.

TECHNICAL SKILLS

Programming & Databases: Python, SQL (MySQL), MongoDB, Core Java

Data Science & ML: Pandas, NumPy, Scikit-learn, Regression, Clustering (K-Means), PCA, Feature Engineering, Model Evaluation

Visualization & BI: Power BI, Matplotlib, Seaborn, Streamlit

Testing & Automation (Prior Experience): Selenium WebDriver, Cucumber BDD, TestNG, Postman

Tools & Platforms: Git, Bitbucket, Jenkins, JIRA

Methodologies: Agile–Scrum, Data Cleaning, EDA, Data Storytelling

DATA SCIENCE PROJECTS

Amazon Music Clustering | Unsupervised Machine Learning

Domain: Music Analytics

- Performed **data exploration, cleaning, normalization, and feature selection** on audio feature datasets. - Built **K-Means clustering models** to group songs based on tempo, energy, danceability, and acoustic features. - Used **Elbow Method and Silhouette Score** to determine optimal cluster count. - Applied **PCA for dimensionality reduction** and visualized clusters to infer genres and moods. - Delivered insights through **data storytelling and cluster interpretation**.

PowerPulse: Household Energy Usage Forecast | Regression Modeling

Domain: Energy & Utilities

- Preprocessed historical household energy consumption data and handled missing

values. - Engineered time-based and usage-based features to improve model performance. - Built and evaluated **regression models** to predict energy consumption. - Analyzed model performance using appropriate **evaluation metrics**. - Provided **actionable insights** to optimize energy usage and demand forecasting.

Ola Ride Insights | SQL, Power BI, Streamlit

Domain: Ride-Sharing & Mobility Analytics

- Cleaned and processed large-scale ride booking datasets. - Wrote complex **SQL queries** to extract operational and customer insights. - Performed **EDA** to identify trends in rides, fares, demand, and driver availability. - Designed an interactive **Power BI dashboard** for business stakeholders. - Built a **Streamlit web application** to present insights in an interactive format.

PhonePe Transaction Insights | FinTech Analytics

Domain: Digital Payments & Finance

- Extracted and analyzed transaction and insurance data using **SQL**. - Identified **top-performing states, districts, and pin codes** based on transaction volume and value. - Created **geo-visualizations and trend dashboards** to analyze user engagement. - Developed a **Streamlit-based analytics app** with clear documentation for stakeholders.

PROFESSIONAL EXPERIENCE

Senior Software Test Engineer

Virtusa Consulting Services Pvt Ltd | Nov 2021 – Present - Validated **backend data using MySQL and MongoDB**, ensuring data accuracy and consistency—directly transferable to data analytics workflows. - Automated UI and API testing using **Selenium, Java, Cucumber, Postman**, improving test coverage and system reliability. - Analyzed system requirements and translated them into **structured test cases and validation logic**. - Collaborated with cross-functional teams to identify data issues and improve overall product quality. - Implemented and maintained **Cucumber BDD frameworks**, enhancing automation efficiency and maintainability.

EDUCATION

Bachelor of Engineering (B.E.) – Metallurgical Engineering
Government College of Engineering, Salem | 2009 – 2013
CGPA: 7.91 / 10

CERTIFICATIONS

- ISTQB Certified Tester – Foundation Level (CTFL)

- ISTQB Certified Agile Tester

ACHIEVEMENTS

- Implemented scalable **BDD automation frameworks**, reducing manual testing effort.
- Streamlined backend **data validation processes** using MongoDB.
- Recipient of **SPOT Award – Certificate of Appreciation** at Virtusa.

ADDITIONAL INFORMATION

Languages: Tamil, English

Career Focus: Data Scientist / Data Analyst / Machine Learning Engineer