

# Rakesh Chaudhary



✉ arcsle09plus@gmail.com


☎ 9998006038/7016731487

🔗 Detailed CV





## 👤 Profile

An accomplished Technology & Research Professional having 6 + Years of experience in all phases of agile software development lifecycle. Passionate about Analytics Engineering Technologies and delivery of effective solutions through creative problem-solving.

## 🎓 Education

**Master of Science - Financial Mathematics**, Maharaja Sayajirao University of Baroda   
Jun 2016 – Jun 2018

## 📜 Certification

- Certified SAFe® 6 Scrum Master 
- Google Cloud Skills 
- Databricks Lakehouse Fundamentals 
- HackerRank 

## 👛 Professional Experience

**Manager - TB Cash Analytics**, Standard Chartered Bank 

May 2024 – present | Bangalore, India

- Contributing as Analytics Engineer to build Cash Data Lake using Bank's various transactional processing systems.
- Leading the Reporting Automations and Cash Data Assets Initiatives aiming to generate Bank's Cash Revenue Opportunities.

**Senior Analyst**, NielsenIQ 

Nov 2018 – May 2024 | Remote

- An Integral Part of Global Innovation and Research Team.

## 🧠 Technical Skills

### Programming Languages

Python | SQL

● ● ● ● ●

### Databases/DWH

MySQL | Oracle | Snowflake

● ● ● ● ●

### Data Transformation Tools

Pyspark | Pandas | Polars | DBT

● ● ● ● ●

### DevOps Tools

Git | Bitbucket | Docker

● ● ● ● ●

### Cloud Services

Databricks | Blob | ADF | S3

● ● ● ● ●

### Project Tracking Tools

JIRA and Confluence

● ● ● ● ●

### Machine Learning Models

Tree-based Models | Azure ML (StackEnsambler)

● ● ● ● ●

### Front End Tools

PowerBI | Tableau | Streamlit

● ● ● ● ●

## 📁 Projects Summary

### Cash Performance View (StanChart)

**Tech: Pyspark, Docker, GIT, JFROG, Shell**

- Architected, Developed and Deployed Reporting Application to monitor the Bank's Cash Revenue (NFI and NII) and Liabilities on Monthly basis.
- Built Analytical Data Models using Bank's Internal OLTP system(Oracle) and data lake(HDFS)
- Automated data extraction using SQL, Pyspark and Shell Scripting.
- Transformed the Data using Python-Polars library to build cost effective data processing modules.
- Discovered the excel workbook level Encryption Method in Python to meet bank compliance.
- 2000 Analytical Sales Level Reports are auto-sent to Bank Sales Executives and auto-shared the 42 Country Level Reports to Country Sales Leadership in Sharepoint Site.

### Stores Data Classification (NielsenIQ)

**Tech: Python, Imblearn, Polars, Pandas, Oracle SQL & Azure ML, Docker, Eliza (ML Ops), JFROG**

- Acting Lead Developer to implement classifier for highly imbalanced binary classification problem.
- Developed and Automated Data Extractions from Oracle OLTP Datababse for 80 countries datasets.
- Performed EDA and explored various techniques to optimized the model performance.
- Reduced total ~50% manual validation efforts(20 FTEs work) by deploying the models for 80 countries.

- Developed the technical POCs for large scale data collection, Ingestion and Validation activities.

#### **Retailer Health Dashboard (NielsenIQ)**

**Tech:** Oracle SQL, ADF, Databricks(Pyspark), Azure-Blob & PowerBI

Used 2 different data sources (oracle database and remote on-prem storage) for data ingestion using Azure Data Factory.

Developed ,deployed & schedule rule-based transformation pipelines using Databricks - Pyspark.

Built store level and item level KPIs for Retailers FMCG Data.

The automated ETL data pipelines processes 20 million records weekly.

#### **Ecom Data Acquisition (NielsenIQ)**

**Tech:** Scrapy, Selenium, RapidFuzz, EasyOCR, Polars, Pandas, Azure Data Lake,Multiprocessing,Streamlit

- Acted as Lead Developer for various mid to large scale webscraping Initiatives.
- Explored and Implemented various robust webscraping techniques like APIs, Sitemaps, Bypass Antibots, Optimal Usage of external proxy services.
- Designed and Developed various Web Scrapers to scrape product images, price, sales, units and categories data (Retailers and Google sites) to extract ~2 million products data every week.
- Developed and deployed various webscrapers for Italy based retailer websites in order to extract the weekly promotional flyers(Avg 20 GB Data Extraction).
- Built various shell script utilities to schedule the webscraper jobs.