



# HARISH ARAVINDH M

## DATA ENGINEER

### OBJECTIVE

To utilize my technical expertise to design and implement efficient data pipelines and architectures. Seeking a challenging role in a dynamic organization to optimize data workflows, improve data quality, and support data-driven decision-making. Committed to delivering scalable and reliable solutions that align with business goals and industry best practices.

### PROFILE SUMMARY

- 2+ years of experience in Analysis, Development and worked for Banking, Financial clients and have experience in various new technologies.
- Deep expertise in PySpark and Python (along Pandas, NumPy libraries) with focus on Data Mutations / Transformations.
- Good Knowledge in PL/SQL with experience in multiple RDBMS.
- Good knowledge and experience in cloud technologies like AWS, Databricks, Azure.
- Excellent analytical, problem-solving and computing skills.
- Strong organizational, planning, inter-personal and communication skills.
- Knowledge of agile software development practices and release management.
- Good Knowledge in SAS programming, worked in migration of SAS to open-source projects, and SAS discovery projects.
- Experience in writing optimized Python programs for reduced runtime and optimized memory usage.
- Worked for various Banking, Financial, Insurance clients in North American regions.

### TECHNOLOGIES WORKED

*Big Data Technologies:* Python, Spark, SAS, Red Hat Linux

*Cloud platforms:* AWS, Azure

*Methodologies:* Agile Methodology

*Database:* MS SQL SERVER, MYSQL, IBM DB2

## TECHNICAL SKILLS

<i>Python:</i>	Python Programming, PySpark, Pandas, NumPy, DS and Algorithms
<i>Spark:</i>	PySpark Dataframes, Spark SQL, Spark Streaming
<i>SAS:</i>	SAS Programming 1: Essentials, Macros, Proc SQL, Data Proc, Proc Freq, Proc Means
<i>Database:</i>	PL/SQL with MS SQL SERVER, MYSQL, IBM DB2
<i>AWS:</i>	S3, EC2, EMR, AWS CLI
<i>Azure:</i>	Azure Fundamentals (AZ900) – Microsoft Certified, Azure SQL Database, Azure Synapse Analytics
<i>Databricks:</i>	Databricks Lakehouse Fundamentals – Academy Accreditation, Notebooks
<i>Other Clouds:</i>	Cognite Data Fusion (CDF Fundamentals)
<i>Languages:</i>	Python, Base SAS, C#, ASP.Net, JAVA, Linux Shell Scripting
<i>Tools:</i>	VS Code, PyCharm, SAS Studio, Visual Studio, JIRA, JTMF, Docker, DB Visualizer, MySQL Workbench, SSMS, SQLDeveloper, MS Office, AWS Console, Azure Portal

## EXPERIENCE

### ***Accenture Solutions Pvt. Ltd.***

Data Engineering Analyst (June 2022 – Till date)  
Data Engineering Associate (March 2021 – May 2022)  
Application Development Associate (December 2020 – March 2021)

## DETAILED WORK EXPERIENCE

### ***Data Engineering Experience in Accenture***

#### ***TD – SAS Remediation*** ***(January 2022 – Till date)***

TD Bank (Toronto-Dominion Bank) is a Canadian multinational banking and financial services corporation. TD Bank decided to migrate from the legacy SAS platform to Python (open-source technologies) to perform their Analytics, Reporting and ETL jobs, which will help them reduce their operational costs.

#### **TD – SAS Remediation Project Overview:**

- Discovery of TD's SAS programs using Accenture's SAS discovery (**Python (Pandas/PySpark) based**) tool to plan and prioritize workload migration.
- Convert around 600 of TD's SAS programs to Python programs.
- Optimize Python programs to match performance of SAS programs

#### **Responsibilities:**

- Worked on creating a detailed inventory of TD's SAS codes, different SAS Procs used across the workloads and data lineage using Accenture's SAS discovery tool to plan and prioritize workload migration.

- Worked on gathering various requirements based on high level discussions with team and detailing fields and layout to extract data from various source systems like IBM DB2, Oracle Database, etc..
- Worked as part of development team for writing new Python programs to replace TD's SAS programs utilizing **Pandas Dataframes and NumPy Arrays** to replace SAS Tables.
- Worked on implementing various code changes for ETL operations.
- Implemented various aggregator functions, data filtering techniques, lookup functionalities using **Pandas** library for data processing.
- Conversions of TD's **ETL, Analytics, and Reporting** SAS programs to **Python (Pandas)** programs, Conversion of repetitive Load and Unload SAS programs into Python with SAS2Py automated workload migration tool.
- Technical checkpoints to enforce Enterprise Data standards and quality.
- Validation and execution of the new python programs.
- Validation of the output datasets, reports against the outputs created by running SAS programs.
- Implementing changes to **AIX** based shell scripts that calls SAS workloads, so the scripts will work on **RHEL** servers to call the new Python workloads.
- Also assisted SIT team for a short period of time after Development is completed.
- Worked on Optimizing the python programs by **implementing multiprocessing using queues, chunking of datasets, using efficient datatypes and vectorizing the code using lambda functions and NumPy library**, formatting the data and writing fixed with files efficiently using NumPy library functions.
- Optimizing the python programs helped in reducing the runtime less than or equal to 2 times the runtime of SAS Programs.
- Worked as part of warranty and post production support team.
- TD is planning to decommission their SAS grid sooner, which will reduce their operating costs.

## ***SYNCHRONY FINANCIALS – SAS Data Modernization*** ***(August 2021 – December 2021)***

Synchrony Financial is a consumer financial services company based of the United States. Synchrony Financial decided to migrate from the legacy SAS platform to Python (open-source technologies) to perform their Analytics, Reporting and ETL jobs, which will help them reduce their operational costs.

### **SYF – SAS Data Modernization Project Overview:**

- Discovery of Synchrony's thousands of SAS programs using Accenture's SAS discovery (**Python (Pandas/PySpark) based**) tool to plan and prioritize workload migration, Identify tables and file systems that can be decommissioned, Create inventory of the SAS grid, and set migration plan and cadence.

### **Responsibilities:**

- Worked on gathering various requirements based on high level discussions with teams and detailing fields and layout to create detailed inventory of Synchrony's SAS programs using Accenture's SAS discovery tools.

- Identify different workloads like **Data wrangling, reporting, analytics**, etc. based on the SAS procs used using our SAS Discovery tool, by reading logs and SAS programs.
- Creating a detailed inventory of **data lineage** for different data warehouses, schemas, tables referenced in SAS Programs using our SAS Discovery tool to help in migration of database to cloud.
- Using the created inventory for identifying simple, complex SAS programs and help in planning migration of workloads, conversion of SAS codes to python.
- Worked on creating a detailed inventory of all the file systems in Synchrony's SAS grid using **Python (Pandas)** and running **Linux system commands to generate a large dataset** for inventory analysis of the SAS grid.
- Worked on identifying various **file formats, SAS tables, datasets like SAS7BDAT, CSV, DAT, TXT in different file systems**, spaces occupied by the datasets, and identifying least accessed to frequently accessed datasets using **Python (Pandas)** codes and mapping them against business units and users to help in planning for migration to cloud and to identify datasets to be decommissioned.
- Now Synchrony Financial is in the initial phase of starting their migration from SAS to Python.

## **ACCENTURE SAS INVENTORY ASSET DEVELOPMENT**

### **INTERNAL (March 2021 – August 2021)**

#### **Project Overview:**

- Development of SAS Inventory tool with Python (Pandas and PySpark Versions) to read clients logs, LSF logs, SAS logs, and SAS programs to create inventory and analyse different workloads, data lineage and mapping business units, users to the workloads, and identifying the complexity of workloads.

#### **Responsibilities:**

- Worked on gathering various requirements based on high level discussions with team and detailing fields and layout to different modules.
- Worked on developing different modules like **readers which read different formats files like log files, LSF logs, SAS logs, SAS programs, SAS programs** inside EGP files to create primary Datasets
- Worked on developing and optimizing **analyzer modules (PySpark and Pandas Versions) which apply transformations** on the primary dataset and create new datasets reading SAS programs and merge the data to create detailed inventory and report on workloads, data lineage and complexity.
- Worked on containerization of the tool with Docker and run tool as docker image.
- Worked on Code building activities, knowledge, and experienced in code versioning and migration to various environment using Bitbucket.
- Worked on **testing and deployment of developed asset** and used it for performing SAS inventory analysis in a real client environment for **Humana Inc.**

***ASE Training Period***  
***(December 2020 – March 2021)***

- Onboarding exam for ASEs conducted by Knowledge Management (LKM) Accenture.
- ASE Stream Training (ASP.NET) conducted by Learning and LKM.
- Accenture essential trainings, Talent Quotient, Delivery Learning Foundation Training (SDLC, Client Data Protection, Agile Methodologies, etc.)
- ASE Stream Training Assessment (ASP.NET) conducted by LKM.

***Other Projects***

- Exporting data from Excel sheets to MYSQL DB using Python.
- Internet Service Provider website using ASP.Net (Capstone Project - ASE Training).
- Enhancement of Dental Pantomograms to detect Jaw Bone Cysts and Tumors using OTSU's Segmentation method (B.E. ECE - Final Year Project).

**Note:** Due to COVID lockdowns in 2020, April 2020 end semester exams (8<sup>th</sup> semester) was postponed to September 2020, so the onboarding date was changed by Accenture, the revised joining date was 4<sup>th</sup> December 2020.

## **EDUCATION**

*Bachelor's Degree:*      ***Electronics and Communication Engineering, (2016 – 2020)***  
Vel Tech Multi Tech Dr.Rangarajan Dr.Sakunthala Engineering College, Avadi.  
Anna University, Chennai.  
CGPA : 8.04/10

*Higher Secondary Schooling:*      Tamil Nadu State Board of Education. (2016)  
92.75%

*Secondary Schooling:*      Tamil Nadu State Board of Education. (2014)  
98.8%

## **CERTIFICATIONS & AWARDS**

### ***CERTIFICATIONS AND ACCREDITATIONS***

- Pursuing AWS Cloud Developer – Associate Certification (DVA-C02)
- Databricks Lakehouse Fundamentals – Academy Accreditation
- Azure Fundamentals AZ900 – Microsoft Certified
- SAS Programming 1 : Essentials Learning – SAS Institute
- Working with Cognite Data Fusion (CDF) Basics – Cognite Academy
- Python Programming – Hackerrank, LinkedIn Skill Assessments

### **AWARDS**

- Received Appreciations and Recognition points from LKM Accenture for completing Techgenics (ASE Training and Assessments) with 90%.
- Received Appreciations and Recognition points from LKM Accenture for Successfully completing DevOps foundation school in Tech Leap.
- Received Appreciations and Recognition points from Manager for impactful delivery for client.
- Got the Outstanding Rating and Promotion from ADL Accenture.

### **CONTACT**

*Mobile:* +91 7395842051; +91 7904009417  
*Email:* [harisharavindhm@gmail.com](mailto:harisharavindhm@gmail.com); harisharavindhm@outlook.com  
*Location:* Ramanathapuram, Tamil Nadu, India. 623503.  
*LinkedIn:* <https://www.linkedin.com/in/harish-aravindh-m-331288129/>  
*Website:* <https://mharavindh.github.io/>