**Vijayasekar Muthumani** 

**Mobile - +91** 9176024472

**Email - vijay2.muthumani@gmail.com**

**Summary:**

* Having 6+ yrs Experience in in both ETL/Big Data Development.
* Having good Domain Knowledge in **Healthcare/ Life Science and HRMS**
* Analysis, Design, Development and Implementation as a **Data Engineer**.
* Hands on experience across **Hadoop** Ecosystem that includes extensive experience in Big Data technologies like**, Spark, Sqoop, Hive, Mongo DB**
* Creating Streaming **Data Pipe Line using AWS** Technologies
* Hands on experience in cloud integration using AWS Services that includes **EMR, S3, Kinesis Firehose, AWS Lambda, AWS Redshift, Glue, Step Function, Athena and Boto3**
* **Hands on Experience Using Snowflake Data Warehouse on AWS Cloud**
* **Load Data From On Premise DB MY SQL to Snowflake Data Warehouse**
* **Experience in Developing Data Visualization dashboard using Tableau Desktop**
* Good Domain Knowledge in HRIS, Healthcare.
* A Quick learner, self-starter, highly motivated individual possessing strong analytical power, and Problem-solving abilities.
* Experience in building large-scale data processing systems and serving as an expert in data warehousing solutions while working with a variety of database technologies.
* Able to integrate state-of-the-art Big Data technologies into the overall architecture and lead a team of developers through the construction, testing and implementation phase.
* Hands on experience in developing, tuning **SQL & PL/SQL for ETL** & data migration involving large data volume.
* Good understanding of Data Warehousing concepts & Dimension modelling.
* Develop new Source and Target mappings and workflow by referring to Mapping/Design docs.
* Good understanding of **ETL Architecture and data warehousing** concepts.
* Good Experience in applying the transformation logic and performance tuning.
* Liaise with onsite counterparts in gathering project requirements and work closely with onshore counterparts in delivering projects.
* Understand project / technical requirements and work with onsite team and Developer to address ambiguities in requirements.
* Work on managing multiple activities/projects.
* Knowledge in NOSQL databases like MongoDB.
* Developed **Spark jobs** and **Hive Jobs** to summarize and **transform data.**
* Experience in working with structured data using **HiveQL,** join operations, **Hive UDFs,** partitions, bucketing and internal/external tables.
* Experience working with **Python, UNIX and shell scripting**.
* Experience in Extraction, Transformation and **Loading (ETL)** of data from multiple sources like Flat files and Databases.
* Experience with complete Software Development Life Cycle **(SDLC)** process which includes Requirement **Gathering, Analysis, Designing, Developing, Testing, Implementing** and Documenting.
* Hands on Experience in Spark architecture and its integrations like **Spark SQL, Data Frames** and **Datasets APIs**.
* Worked on Spark for enhancing the executions of current processing in Hadoop utilizing Spark **Context, Spark SQL**, **Data Frames**
* Involved in converting **Hive/SQL** queries into Spark transformations using  **Spark SQL** and **Python.**
* Hands on experience Using **Hive Tables by Spark**, performing transformations and Creating Data Frames on **Hive tables** using Spark.
* Exposure to **CI/CD tools**
* Worked with **waterfall and Agile** methodologies.
* Mentor junior team members and help to solve their technical/functional queries.
* Experienced working towards building a strong team/work environment, and have the ability to acclimate to new technologies and situations with ease.
* Highly motivated team player with excellent analytical, problem solving, interpersonal and communication skills.
* Interfacing and interacting with the Technical teams, Business analyst and End Users.
* Ability to communicate verbally and via written documentation and work under minimal supervision.

**TECHNICAL SKILLS:**

Cloud Computing: AWS

Big Data Technologies: Apache Spark, Hadoop

Spark: Spark Core, SparkSQL

Interactive Data Analysis: Databricks , Snowflake

Data Warehouse: RedShift

Programming Languages: Python

Databases: SQL, Oracle, ODI, MySQL, PostgreSQL

Scripting: Shell Scripting, HTML

AWS Services: EMR, EC2, S3, Lambda, IAM, CloudFormation, SNS, CloudWatch

Schedulers: Airflow

Version Control: GitHub

CI/CD: Jenkins

Containers: Docker

Project Management Tools: Jira, SharePoint, Confluence

ETL Tools: ODI

**AREAS OF EXPERTISE**

* End to End data pipeline design, development, testing and deployment
* Interactive **Data Analysis** using **Databricks, Snowflake** and AWS
* Big Data Processing using **Apache Spark**
* Develop and implement **Bigdata processing** pipelines on **AWS**
* Python Programming for **PySpark and AWS services**
* Ability to quickly master new concepts.
* Bigdata workflow management tools migration to Airflow

**PROFESSIONAL CERTIFICATIONS**

* AWS Certified Data Analytics – Specialty (AWS DAS C01)
* Certified Tableau Speciality

**EDUCATION**

* M.Sc Software Engineering, Annamalai Uiversity (2007 - 2009), Chennai, India

**PROJECT EXPERIENCE**

**Tango Analytics, Texas Jan 2023 to till date**

**Sr Data Engineer**

* Analysing large amounts of data sets
* Designed and Implemented **Big Data Analytics architecture/pipeline.**
* **Load Data from MySQL DB to S3 and to AWS Redshift using AWS Lambda**
* Developing AWS Data pipeline using **Glue**
* Developing Python code using **Pyspark** to Processing Data in **Spark**.

Experienced in fact dimensional modelling (Star schema), and SCD (Slowly changing dimension)

* Took part in the meetings with Sr. Business Analyst and managers and understood various process plans, business process, and functionality in detail.
* Manage daily activity to meet all deadlines.

**Evironments :- MySQL, AWS S3, AWS Redshift, Glue, Lambda, Python, PySPark**

**Lnt Infotech Jan 2022 to Jan 2023**

**Sr Specialist - Data Engineering (Lead)**

**Description: -**Enabled the marketing and advertising teams for content customization, targeted and personalized advertising across different platforms by developing a **data management** platform, by transforming and organizing the analytical and clickstream data collected from different data sources. Developed data pipelines using **PySpark and SparkSQL** and implemented them in the AWS cloud. **Developed Stable**, Resilient and **Cost-effective Data Pipelines** in the **AWS Cloud. Created IAM** roles, **S3 buckets**, Bucket policies, Security Groups, **Cloud Formation Templates** and other AWS infrastructure required. Developed **Python Scripts** to launch **AWS** infrastructures and terminate them once the pipeline is completed. Perform Ad-hoc analysis and debugging using DataBricks, Snowflake. Performed Code review for **applications & frameworks** developed by the team. Migrated the Big Data pipeline workflows to **Apache Airflow**.

**Roles & Responsibilities:**

* Built a POC wherein we design an ingestion Framework would ingest the files from **SFTP to HDFS** using **Apache Scoop and Hive** and Ingest Financial Data into HDFS
* Built **Validation Framework** that would compare the data at source **RDBMS** Systems and Destination Hive tables.
* Analysing large amounts of data sets to determine optimal way to aggregate and report on these data sets.
* Designed and Implemented **Big Data Analytics architecture/pipeline.**
* **Load Data From On Premise DB MY SQL to Snowflake Data Warehouse and Performing Data Manipulation using Snow SQL**
* Experienced in writing **Spark Applications Python**.
* Built multiple notebooks and piped all of the jobs together, built an **ETL pipeline** to making wear algorithm predictions of the model, and writing the outputs to **AWS Redshift**.
* Analysed the **SQL scripts** and designed the solution to implement using Spark.
* Developed custom aggregate functions using **Spark SQL** and performed interactive querying.
* Implemented ADB job scheduler for authoring, scheduling and monitoring Data Pipelines
* Developed Python code to gather the data from **API’s** and designs the solution to implement using **Spark.**
* Implemented Kinesis **model** which pulls the latest records into S3 Bucket.
* Imported data into **ADLS** from various API’s and **SQL databases** and files using **ETL** and from streaming systems.
* Loaded all data-sets into Hive from Source CSV files using spark from Source CSV files using Spark
* Exported the analysed data to Teradata using Sqoop for visualization and to generate reports for the BI team.
* Migrated the computational code in **HQL to PySpark.**
* Completed data extraction, aggregation and analysis in **HDFS by using PySpark** and store the data needed to **Hive.**
* Experienced in fact dimensional modelling (**Star schema**), transactional modelling and **SCD** (Slowly changing dimension)
* Working on the integration of data engineering and Tire Wear model with data science team, to make the correct output predictions and working on fixing other issues.
* Sound knowledge in programming Spark using Scala.
* Extract Transform and Load data from Sources Systems to **AWS RDS** services using a combination of , T-SQL, Spark SQL . Data Ingestion to one or more AWS Services - (AWS RDS, AWS Redshift ) and processing the data in In **AWS Snowflake.**
* Experienced in Importing and exporting data into **HDFS** and Hive using Sqoop.
* Deploying Spark jobs in **Amazon EMR** and running the job on **AWS clusters**.
* Strong Experience in implementing Data warehouse solutions in **Amazon web services** (AWS) Worked on various projects to migrate data from on premise databases to AWS RDS and S3.
* Was Responsible in Managing tasks and deadlines for Offshore ETL teams
* Interacting with the customer/Business Analyst on requirement gathering
* Involved in detailed technical Design

**Environment:** AWS, Spark, Python, Snowflake, Jenkins, Airflow, Unix

**CMA-CGM** **Aug 2020 to Sep 2021**

**Sr AWS Big Data Engineer**

**Roles & Responsibilities:** -

* Identify, design, and implement batch and realtime data pipeline using AWS Big Data Technologies.
* Developed pipelines to pull data from Redshift and send it to downstream Databases like PostgreSQL.
* Importing data from files from S3, and from SQLWorkbench data pumper to Redshift tables.
* Experience building distributed high-performance systems using Spark, python and Airflow for Database migrations from Traditional Data Warehouses to Spark Clusters.
* Use Airflow to define workflow.
* Extensively involved in writing SQL queries (Subqueries and Join conditions), PL/SQL programming.
* Data visualization using Tableau and accessed aggregations using SQL Clients PostgreSQL & SQL-Workbench.
* Participating in Daily Scrum meetings.
* Took part in the meetings with Sr. Business Analyst and managers and understood various process plans, business process, and functionality in detail.
* Manage daily activity to meet all deadlines.
* Implemented and operated **on-premises Hadoop clusters** from the hardware to the application layer including compute and storage.
* Uploaded and processed more than 30 terabytes of data from various structured and unstructured sources into **HDFS (AWS cloud)** using Sqoop and Flume.
* Handled importing other enterprise data from different data sources into **HDFS** using Sqoop and performing transformations using **Hive**, **MapReduce** and then loading data into HBase tables.
* Involved in Cluster Maintenance and removal of nodes using Cloudera Manager.
* Collaborated with application development teams to provide operational support, platform expansion, and upgrades for **Hadoop** Infrastructure including upgrades to **CDH3**.

**Environment:** Hadoop, Pyspark, Talend Studio, Amazon S3, Redshift, My SQL, PostgreSQL, Tableau.

**BCT Consulting (Client – AstraZeneca) Oct 2018 to May 2020**

**Lead Data Engineer**

**Roles & Responsibilities: -**

* Developed Spark Applications by using Pyspark and Implemented Apache Spark data processing Project to handle data from various RDBMS and Streaming sources.
* Worked with the Spark for improving performance and optimization of the existing algorithms in Hadoop
* Using Spark Context, Spark-SQL, Data Frame
* Composer orchestration using Python.
* Used Spark Streaming APIs to perform transformations and actions on the fly for building common Learner data model which gets the data from Kafka in near real time and persist it to Cassandra.
* Developed Kafka consumer API in Scala for consuming data from Kenisis topics.
* Consumed XML messages using Kafka and processed the xml file using Spark Streaming to capture UI updates.
* Developed Preprocessing job using Spark Data frames to flatten JSON documents to flat file.
* Load D-Stream data into Spark RDD and do in memory data Computation to generate Output Response.
* Developed data integration modules using AWS Step Functions, AWS Glue and Python modules.
* Migrated an existing on-premises application to AWS.
* Used AWS services like EC2 and S3 for small Data sets processing and storage.
* Experienced in Maintaining the Hadoop cluster on AWS EMR.
* Imported data from AWS S3 into Spark SQL, Performed transformations and actions on SQL.
* Good understanding of Cassandra architecture, replication strategy, gossip, snitches etc.
* Worked from Scratch in Configurations of AWS Kinesis such as Mangers and Brokers
* Experienced in creating data-models for Clients transactional logs, analyzed the data from Casandra
* Tables for quick searching, sorting and grouping using the Cassandra Query Language (CQL).
* Tested the cluster Performance using Cassandra-stress tool to measure and improve the Read/Writes.
* Stored in Hive to perform data analysis to meet the business specification logic.
* Used Apache Kafka to aggregate web log data from multiple servers and make them available in
* Downstream systems for Data analysis and engineering type of roles.

**Environment:** Pyspark, Redshift, AWS S3, PostgreSQL, Kinesis, Airflow.

**Wipro (Client - Khols) Mar 2018 to Oct 2018**

**Test Engineer**

**Responsibilities:**

* Identifying Testing Objective.
* Analysis of test requirement.
* Test case documentation.
* Identifying the ETL Process and validating the Transformation Logic for ETL Testing.
* Testing Snapshot and Incremental Load in ETL Testing.
* Writing complex SQL Queries using Analytical function for Data Analysis from Source to Target.
* Identify the Test Requirements inbound to Hadoop – Big Query mapping and Data Analysis
* Performing Hadoop Testing Using Hive, Big Query (Google Cloud Platform)
* Performing DB to File and File to DB data validation
* Ensure all the QA Activities are as per schedule.
* Manage daily activity to meet all deadlines.
* Prepare Test Planning, Test Strategy and Test Approach Document.
* Provide Post Implementation Support for UAT for Preparing Test Data and Assisting Business users.
* Preparation of Test Scenarios & Cases, Expected Results for the release.
* Provide Input on client clarification.
* Conduct Daily Scrum Meeting internally within Testing Team.
* Providing feedback/updates to relevant Project Managers.
* Participation in Daily Scrum Stand-Up meeting along with the managers, Developers.
* Conduct Scrum Meeting for internal Testing team.
* Facilitate Scrum Retrospective for internal Testing team along with Test Manager.
* Took part in the meetings with Sr. Business Analyst and managers and understood various process plans, business process, and functionality in detail.

**Environment:** Hive, Scoop, GCP, Python, SQL Server.

**Manpower (Client – IQVIA) Oct 2017 to Mar 2018**

**Sr System Engineer**

Identifying Testing Objective and Validation Scope.

• Analysis of test requirement

• Test case documentation

• Identifying the ETL Process and validating the Transformation Logic for ETL Testing.

• Testing Snapshot and Incremental Load in ETL Testing.

• Writing complex SQL Queries using Analytical function for Data Analysis from Source

to Target.

• Identify the Test Requirements inbound to Hadoop mapping and Data Analysis

• Performing Hadoop Testing Using Hive, Impala, Scoop and Spark

Ensure all the QA Activities are as per schedule.

• Manage daily activity to meet all deadlines.

• Prepare Test Planning, Test Strategy and Test Approach Document.

• Provide Post Implementation Support for UAT for Preparing Test Data and Assisting

Business users

• Preparation of Test Scenarios &amp; Cases, Expected Results for the release

• Provide Input on client clarification.

• Conduct Daily Scrum Meeting internally within Testing Team

• Providing feedback/updates to relevant Project Managers.

• Participation in Daily Scrum Stand-Up meeting along with the managers, Developers

• Conduct Scrum Meeting for internal Testing team.

• Facilitate Scrum Retrospective for internal Testing team along with Test Manager

• Took part in the meetings with Sr. Business Analyst and managers and understood

various process plans, business process, and functionality in detail.

• Raising defects where discrepancy found in the database and reporting it to the onshore

development team.

• Was also involved in Hive Development.

• Creating Hive tables and working them on HiveQL

• Imported data using Scoop to Load data from MYSQL to HDFS using Scoop Statements

**Environment:** Hive, Scoop, Impala.

**Icon Clinical Research May 2013 to May 2017**

**Sr Test Analyst**

**Project Perspective**

The Enterprise Data Warehouse (EDW) is aimed at extracting the clinical trials data from

Multiple disparate source systems in a meaningful way to CDR and OMR repositories by

building data warehouse and data marts to meet the overall objectives of the ICONIK (ICON

Immediate Knowledge) initiative, with the objective of providing secure access to the clinical

data with appropriate analysis tools in an operational, holistic and aggregate manner that

supports their decision making

**•** Identifying Testing Objective and Validation Scope.

• Analysis of test requirement

• Test case documentation

• Identifying the ETL Process and validating the Transformation Logic for ETL Testing.

• Testing Snapshot and Incremental Load in ETL Testing.

• Writing complex SQL Queries using Analytical function for Data Analysis from Source

to Target.

• Identify the Test Requirements inbound to Hadoop mapping and Data Analysis

• Performing Hadoop Testing Using Hive and Impala

• Verifying data in HDFS ecosystem using Cloudera HIVE

• Monitoring Hadoop Jobs using OOZIE Dashboard

• Middleware API testing (XML and JSON) using SOAPUI.

• Performing BI Reporting testing.

• Ensure all the QA Activities are as per schedule.

• Manage daily activity to meet all deadlines.

• Prepare Test Planning, Test Strategy and Test Approach Document.

• Provide Post Implementation Support for UAT for Preparing Test Data and Assisting

Business users

• Preparation of Test Scenarios &amp; Cases, Expected Results for the release

• Provide Input on client clarification.

• Conduct Daily Scrum Meeting internally within Testing Team

• Providing feedback/updates to relevant Project Managers.

• Participation in Daily Scrum Stand-Up meeting along with the managers, Developers

• Conduct Scrum Meeting for internal Testing team.

• Facilitate Scrum Retrospective for internal Testing team along with Test Manager

• Took part in the meetings with Sr. Business Analyst and managers and understood

various process plans, business process, and functionality in detail.

• Raising defects where discrepancy found in the database and reporting it to the onshore

development team.

**Environment:** ODI, OBIEE, XML, JSON , Oracle.Cloudera

**Helios & Matheson (Client – Bank of America) July 2012 to Jan 2013**

**Sr Test Analyst**

**Project Perspective**

Payments coming for Indian Retail banking customer will be processed via SWIFT messages and

XML in core banking system in Bank of America and the same will be validated and responded back

• Reviewed BRD and LLD/HLD to identify business process and design.

• Identifying the Objective and Testing Scope.

• Prepared Test plan for the project based on the BRD and LLD/HLD documents

• Ensure all the QA Activities are as per schedule.

• Validating SWIFT Messages flat files against the XML

• API Testing using SOAP UI

• SPOC for all QA related Queries and deliverables.

• Manage daily activity to meet all deadlines.

• Preparation of Test Scenarios &amp; Cases, Expected Results for the release

• Involved in End to End testing.

• Providing feedback/updates to relevant Project Managers.

• Participation in Daily Stand-Up meeting along with the managers, Developers

• Took part in the meetings with Sr. Business Analyst and managers and understood

various process plans, business process, and functionality in detail.

**Environment:** ODI, OBIEE, XML, JSON , Oracle.

**Ciber (Client – Ford Motors) Aug 2011 to Apr 2012**

**Consultant**

Analyzing requirements and understanding business process.

• Thorough understanding of reengineering and designing processes.

• Doing reviews of standard operating procedures and updating the feedback to the

Business Analyst

• Involved in planning and estimation meeting.

• Preparation of Usecases, Test Scenarios (Basic path and Alternate Path), &amp; Cases, Test

Matrix, Expected Results for each release

• Involved in basic flow and Alternate flow (End to End) testing

• Preparation of Test Bed to enable QA Team to effectively Test all Functional Scenarios.

• Providing feedback/updates to relevant Project Managers.

• Participation in Daily Stand-Up meeting along with the managers, Developers

• Took part in the meetings with Sr. Business Analyst and managers and understood

various process plans, business process, and functionality in detail

**Mastek (Client – Fidelity Investments) Aug 2010 to May 2011**

**Sr Test Engineer**

Analyzing requirements.

• Doing reviews of requirements and updating the feedback to the Managers

• Preparing QA Metrics and tracking the QA status as per the plan.

• Preparation of Master Test Scenarios &amp; Cases, Expected Results for each release

• Involved in End to End functional testing

• Involved in Regression Testing at SAT

• Preparation of Test Bed to enable QA Team to effectively Test all Functional Scenarios.

• Providing feedback/updates to relevant Project Managers.

• Responsible to manage the requirement throughout the project lifecycle like reviewing

regularly and changing the necessary Test cases, Scenarios and Test Data’s accordingly

• Scheduling, Managing and Tracking the testing Activities as per the project plan

• Participation in Daily Stand-Up meeting along with the managers, Developers

• Took part in the meetings with Sr. Business Analyst and managers and

understood various process plans, business process, and functionality in detail

**Icomm Technologies Nov 2005 to July 2010**

**QA Engineer**

**Product Name – formulaH**

**Domain – Human Resource Management System, Payroll**

**Platform/ Language – Windows/Asp, C#.Net 2.0, ASP.Net 2.0, MS SQL Server 2005, IIS**

**Designation – Engineer – Quality Assurance**

**Duration – November 2005 to Jul 2010**

**Product Perspective – formulaHR**

formulaHR is the only truly integrated HR enterprise solution (SaaS) that delivers the framework

for effective human resource management and growth, binding people and complex HR processes, while

aligning the workforce with corporate objectives

formulaHR is an Enterprise HR Application tailored to various Verticals such as BPO, ITES,

Cement Industries, Cosmetic Industries, Manufacturing and Process Industries etc. This product brings

an end to end paperless solution from Employee daily Attendance to Payroll. This product has brought

tremendous Cost Savings to the above mentioned Projects. The various Modules in formulaHR are HR

Info System, Org Structure, e-Recruitment, Attendance Management System, Leave Management System,

Payroll, Training, Learning &amp; Development, Performance Management System, Employee Diary,

Employee Relationship, Work Place Communications, Help Desk, etc. Client Specific Modules are

provided accordingly.

Responsibilities as QA Engineer

 Preparation of Master Test Scenarios &amp; Cases for each Client/Module.

 Involved in End to End functional testing for six formulaHR rollouts.

 Involved in Regression Testing at alpha stage for six formulaHR rollouts.

 Preparation of Module wise User Manuals which include a diagrammatic description of the Work

Flow of the module.

 Preparation of Master Test Data for Client Demonstration &amp; Implementation covering all modules.

 Involved in Data Mapping, Data Migration Testing using Source and Target Files

 Preparation of Test Bed to enable QA Team to effectively Test all Functional Scenarios.

 Maintaining Documents which contain the list of unfixed Bugs in various versions.

 Gathering and analyzing requirements.

 Doing database testing to ensure the Data Manipulation in the Backend.

 Provide business analysis throughout all phases of project lifecycles and documenting solutions.

 Doing reviews of requirements and designs with stakeholders.

 Maintaining application issue logs to capture business process system issues.

 Writing test scenarios and testing new applications and processes.

 Providing feedback/updates to relevant Project Managers.

 Building and maintaining solid relationships with Technical teams and end Users.

 Keeping up to date documentation.

 Support of acceptance testing phases.

 Delivering systems training to end-users.

 Acted as SPOC for Issue resolution from both business and technology ends.

 Planning, Tracking and Reporting to all the stake holders

 Was a part of 24x7 support team for end-users in Post implementation phase

 Successful End to End Client Demonstration &amp; Implementation for the following projects.

o CSS Corp - formulaHR Implementation &amp; Support

o 247Customer - formulaHR Implementation &amp; Support [http://www.247people.in]

o Deloitte – e-recruitment Implementation &amp; Support [http://dpeople.co.in]

o Agility Logistics - formulaHR Implementation &amp; Support [http://formulahr-agility.com]

o Ashok Pirmal Group - formulaHR Implementation [http://formulahr-apg.com]

o UltraTech - Cements - formulaHR Implementation

o Tata Projects Limited - Custom HR Implementation [http://tplhronline.com/]

o Concorde Motors - Custom HR Implementation [http://concordemotors-formulahr.com/]

o Tata Capital Limited - Custom HR Implementation [http://202.71.144.251]

o Velankani Infrastructures - formulaHR Implementation [http://formulahr-velankani.com]

o Birla Carbon - Employee Performance Diary [http://birlacarbonconnect.com]]

o Grasim Cements - Custom HR Implementation [http://www.grasimsampark.com]

o INX India Limited - Online Recruitment System

o Cholamandalam - e-recruitment Implementation &amp; Support [http://cholapeople.co.in]

o Indian Oil Corporation - Rewards Management System [http://www.iocxtrapower.com]

o Renaissance - Online R&amp;D Data Servives for Nerolac

o Grasim Cements - Cement Tracking System - Inventory &amp; Goods Management System

[http://www.grasimcemtrack.com]

o Raymond - formulaHR Implementation [http://formulahr-raymond.com]

o 247 Customer Guatemala - formulaHR Implementation [http://gua.247people.in]

 Induction Programs on HR Verticals were given to New Joiner’s.

1.How conversant are you with Python/Scala on a scale 1-5

Ans: 3

2. Would you be interested to take a data structures and Algorithm test in any programming language of your choice

Ans: Yes

3. What Development Techniques do you adopt for your data project?

* ETL Development worked on ODI
* AWS: Redshift, AWS Glue, AWS SNR,
* Lambda Function Step Function-implement the workflow and everything.
* Also Pretty much knowledge AWS in CloudFormation for any implementation

1. Explain in brief what measures were taken to assure data quality.

Ans: Firstly, Need to understand Business:

* What kind of data we are handling what is the business behind that.
* Need to understand core fields and data actually using in data warehouse or pulling the data
* ETL testing

5.What data Modelling approach did u adopt for enabling Business Intelligence If any

Ans:

* Star Schemas
* dimension model: acp1 gcp2

6.What technical gaps do you want to work on once you join our organisation?

Ans: 1st would like the understand the business delive r the quality product and then could work as Subject matter expert , lead the Team