**Rashid Mohammed**

**Sr Data Engineer** +1 614 592 3452

# PROFESSIONAL SUMMARY



* Over **Eleven years** of extensive experience in complete Software development life cycle (**SDLC**) with a strong background in Cloud Data Engineering platforms **GCP** and **AWS**, ETL development and implementation of various data warehouses/data marts.
* **AWS Certified** Data Engineer with hands on experience in services like **S3**, **CloudWatch**, **Redshift**, **Lambda**, **Glue** used for Data migration and data Retrieval.
* **GCP** Data Engineering hands on experience in a project for Data migration using services like **cloud storage**, **Composer**, **Bigquery**, **Airflow**, **Dataflow**, **PubSub**, **looker** data models.
* Python Beginner experience in writing cloud functions, importing libraries to do data slicing, data retrieval.
* Strong understanding and experience in Azure cloud services including Azure Data Lake, Azure Data Factory, Azure SQL Data Warehouse, Azure Synapse and Azure Cosmos DB, Data Bricks.
* Consulted for multiple industry domains Auto and Home **Insurance**, **Retail** Supply Chain, **Education**, **Banking** and **Finance** with familiarity to data from these domains.
* Extensive experience as ETL Developer moving data from all different sources and targets using **Informatica Power Center 9.6.1,9.5, 8.x, 7.x.**
* Worked with **Informatica** **Data Quality 9.6.1/10.1/10.4** (IDQ) toolkit, Analysis, data cleansing, data matching, data conversion.
* Demonstrated experience in Change Data Capture using **Informatica CDC**, **SQL Server CDC** and **Date control** techniques.
* Good Understanding & knowledge of Ralph Kimball & Bill Inman Methodologies.
* Extensively worked on **Dimensional modeling** (Star/Snow flake), Data migration, Slowly Changing Dimensions, Data cleansing and Data Staging of operational sources using ETL processes and providing data mining features for data warehouses.
* Experience in integration of various Operational Data Source (**ODS**) with Multiple Relational Databases like Oracle pl/sql, SQL Server, and DB2 and Worked on integrating data from flat files like fixed width and delimited.
* Actively involved in Performance SQL Tuning, **ETL tuning** and **Error handling**.
* Experience in implementing the **complex business rules** by creating transformations, re-usable transformations (Expression, Aggregator, Filter, Connected and Unconnected Lookup’s, Router, Rank, Joiner, Update Strategy), and developing complex Mapplets and Mappings.
* Experience in Data Modeling (Logical and Physical Design for distributed databases, Reverse-engineering and Forward-engineering using Erwin).
* Extensive experience working closely with **BI team** to complete the dashboard requirements on multiple tools including **Microsoft Power BI, Cognos, Tableau, Business objects, OBIEE**.
* Excellent knowledge of studying the data dependencies using Metadata stored in the Informatica Repository and preparing batches for the existing sessions for scheduling of multiple sessions.
* Proficient in understanding business processes/requirements and translating them into technical requirements.

Experience in building Data Warehouse from different **OLTP** systems.

* Experience with Teradata as the target for the data marts and worked with BTEQ, Fast Load, and Multi Load.
* Worked on **Exception Handling** Mappings for Data Quality, Data Profiling, Data Cleansing, Data Validation
* Experience in shell scripting to run ETL mappings and file handling.
* Developed Test Scripts, Test Cases, and SQL QA scripts to perform Unit testing, System Testing and Load testing.
* Seamlessly migrating the Code from Development to Testing, UAT and Production.
* Expertise over preparing report specifications, database designs to support the reporting requirements
* Perform review of Session log files to trace causes of bottlenecks.

# EDUCATION

* Bachelor of Science (ECE).

Jawaharlal Nehru Technological University Hyderabad, India

* Masters in Electrical Engineering (MSEE) Northwestern Polytechnic University, CA

# TECHNICAL EXPERIENCE

|  |  |
| --- | --- |
| **ETL Tools:** | Informatica Power Center, Power Connect & Power Exchange, Informatica Data Quality, Informatica CDC |
| **Cloud Technology:** | **AWS Certified Cloud Practitioner, GCP Certified** |
| **Data Modeling:** | ERwin 7.3/4.0, UML Rational Rose, Visio 2007 |
| **Databases:** | Oracle pl/sql, TeraData, DB2, SQL Server 2018, DB2 |
| **Reporting Tools:** | Cognos, Business Objects, Tableau, Microsoft Power BI |

# Scripting Languages: Unix Shell Scripting

**SQL Tools:** Toad, Rapid SQL, MS Management Studio, Oracle pl/sql designer,SQL\*PLUS,WinSQL,Teradata Studio.

# PROFESSIONAL EXPERIENCE

# American Family Insurance Mar 2021 – Till Date Sr Data Engineer

**Responsibilities:**

* Build data pipelines in airflow in **GCP** for ETL related jobs using different airflow operators.
* Experience in Google Cloud platform (**GCP**) Dataproc, GCS, Cloud functions, BigQuery.
* Used cloud shell SDK in Google Cloud platform (**GCP**) to configure the services Data Proc, Storage, BigQuery.
* Hands on experience is GCP services like GCP **Dataproc**, **Pubsub**,**Data flow** & **firestore**.
* Building and deploying CI/CD pipelines using Terraform and Jenkins.
* Experience & Certified in AWS cloud platform

Used services like AWS S3,AWS **Redshift**, & **Lambda** , **RDS** for data migration from on prem to cloud.

* Develop Cloud Functions in Python to process JSON files from source and load the files to BigQuery.
* Experience in developing code in GITLAB, Visual Studio, Cloning of code, forking code to branches.
* Worked as ETL developer in development of an **Enterprise Information Model**
* Development of **Informatica CDC** between heterogeneous source & Target when staging data in

Acquisition layer.

* Worked with **Informatica** **Data Quality** 10.4.1 (IDQ) toolkit, Analysis, data cleansing, data matching, data conversion
* Experience Working with **Informatica** **Power center/IDQ** transformations.
* Experience doing source analysis by Profiling using **IDQ business analyst** & building score cards to monitor refined data phases.
* Having experience in development, maintenance, support and enhancement in

**Informatica Power Center**, **Teradata**, Informatica **Data Quality** (IDQ), **Informatica Big Data Management** (BDM), Oracle pl/sql, **GCP**, **AWS**

* Handling of CTE’s, subqueries, functions, stored procedures & triggers in oracle pl/sql and sql.
* Development of UNIX shell scripts and for automation of ETL processes.
* Interacted with the Business Users to analyze the business requirements and transform the business requirements into the technical requirements.
* Follow the quality standards and guidelines for configuration management, testing, project planning and execution.
* Developed the test scripts using SQL for the unit testing.
* Presented curated data ready for processing power BI, Cognos, GCP looker.

# Kroger Technologies, Cincinnati OH Dec 2018 – Mar 2021 Sr IDQ Developer

**Responsibilities:**

* Worked as a **Sr IDQ Developer** performing data quality and implementing continuous Data Governance for many different applications like pricing, promotion, customer, HR etc..
* Implementation of tailored data quality model to bounce Critical Data Elements (CDE’s) through the model to get the refined data.
* Worked on creating Profiles, PDO, LDO and Address Doc shared objects
* Created Reports, Scorecards and Dashboards
* Created IDQ profiles, mapplets, rules, and scorecards for improvement of data sourced from more than 60 local ERPs before loading to Global MDM system.
* Designed/created IDQ reference tables to automate global business rule population to MDM system.
* Design and Implementation of Exception Management framework using Informatica version 10 developer and Analyst tool.
* Developing the Business Initiatives and Business Glossaries in IDQ - Analyst UI.
* Designing and developing the mapping specification and Rule specification to convert the business rules in Informatica mapplets.
* Implementation of **CDC** using the date control techniques to bring current data to avoid reduplicating of work.
* Performed data quality on **streaming** data for few applications using **IDQ Applications (10.4)** and Java API’s.

Providing **remediation** data to application teams to perform manual remediation and also do automated remediation using **exception** handling, selenium tools as per application.

* Migration of application data to Microsoft Data lakes **(ONPREM to OFFPREM).**
* Doing Data Quality on OFFPREM using **Microsoft Azure, spark, Notebooks** and customized python scripts and experience in Azure MS SQL server.
* Involved in implementing the Land Process of loading the customer Data Set into Informatica **MDM** from various source systems.
* Participated in team meeting with Business Analysts of other teams for requirement gathering and transforming it into technical requirements.
* Performed Unit Testing, System Testing and Regression testing and validated the data.
* Involved in testing Informatica mappings using the test data scenarios and upgraded the QA Test Plans.
* Involved in Fine tuning the Informatica Code (mappings and Sessions), Stored Procedures and

SQL queries to obtain optimal performance and throughput

* Defect Fixing & running daily workflows to monitor redone changes in all environments.
* Creating build documentation used for migration of codes & mappings from one environment to another.
* Building of power BI cubes, Looker models and cognos models for reporting and analysis.
* Extensively used almost all the transformations, which includes (Sequence Generator, Expression, Filter, Router, Sorter, Rank, Aggregator, LOOKUP (Static and Dynamic), Update Strategy, Source Qualifier, Stored Procedure, Joiner, Normalizer, XML Source Qualifier).
* Extensively used all Power Center/Power mart capabilities such Target override, Connected, Unconnected and Persistent lookup’s.
* Worked and responsible for running the CA7 Jobs.

# Hastings Mutual Insurance, Hastings MI Mar 2018 – Dec 2018 IDQ Developer

**Responsibilities:**

* Worked as IDQ Lead developer in development of an **Operational Data Model** where users can write off reports.
* Designed and Developed IDQ plans with source system specific requirements using Rule based Analyzer, Weight based Analyzer, Address Validator, Bigram etc.
* Attended team meetings as an ETL/Informatica technical advisory to offer opinion and guidance to managers to select the best plan of approach at project startup meetings.
* Designed and developed mappings with optimal performance using Aggregator, Joiner,

Normalizer, Rank, Sequence Generator, Un-Cached & various kinds of Cached Lookup, Connected-Unconnected-Source Target pre-and post-load Stored Procedure Transformations, Update Strategy, and Union.

* Implementation of **Informatica CDC** between heterogeneous source & Target when staging data in Acquisition layer (**Operational Data Store).**
* Implementation of SQL Server CDC between homogenous source & target when moving data into Architected layer (**Operational Data Store).**
* Participated in team meeting with Business Analysts of other teams for requirement gathering and transforming it into technical requirements.
* Performed Unit Testing, System Testing and Regression testing and validated the data.
* Involved in testing Informatica mappings using the test data scenarios and upgraded the QA Test Plans.
* Used database objects like views, partitioning for accomplishing complex logical situations and implemented Type II Slowly Changing Dimensions.

Involved in Fine tuning the Informatica Code (mappings and Sessions), Stored Procedures and SQL queries to obtain optimal performance and throughput

* Defect Fixing & running daily workflows to monitor redone changes in all environments.
* Creating build documentation used for migration of codes & mappings from one environment to another.
* Extensively used almost all the transformations, which includes (Sequence Generator, Expression, Filter, Router, Sorter, Rank, Aggregator, LOOKUP (Static and Dynamic), Update Strategy, Source Qualifier, Stored Procedure, Joiner, Normalizer, XML Source Qualifier).
* Experience in using **SSIS** tools like Import and Export Wizard, Package Installation, and **SSIS Package Designer**.
* Experience in **importing/exporting** data between different sources like Oracle pl/SQL, /Access/Excel etc. using **SSIS/DTS** utility.
* Expertise in using **global variables**, **expressions** and **functions** for the reports with immense experience in handling sub reports in **SSRS**.
* Extensively used all Power Center/Power mart capabilities such Target override, Connected, Unconnected and Persistent lookup’s.
* Design and development of data quality processes including parsing, analysis, standardization, and enrichment at point of entry and batch modes; Deploy mappings that will run in a scheduled, batch, or real-time environment.
* Designed and implemented best practices for Informatica repository usage, security controls, version control, code migration, and scripting, logging, and other development and deployment procedures.
* Worked and responsible for running the CA7 Jobs.

# Alliance Data, Columbus OH Aug 2017 – Mar 2018 Sr Data Quality Analyst

Alliance Data Card Services is a provider of integrated credit and marketing services including private label, co-branded and commercial credit card programs. Alliance Data Card Services manages over 135 card programs for retail brands.

**Responsibilities:**

* Worked as ETL developer in development of an **Enterprise Information Model**
* Development of **Informatica CDC** between heterogeneous source & Target when staging data in

Acquisition layer.

* Worked with **Informatica** **Data Quality** 9.6.1/10.1 (IDQ) toolkit, Analysis, data cleansing, data matching, data conversion
* Experience Working with **Informatica** **Data Quality** 10.1 transformations.
* Experience doing source analysis by Profiling using **IDQ business analyst** & building score cards to monitor refined data phases.
* Involved with Informatica team members in Designing, document and configure the **Informatica IDQ suite** to support loading, cleansing, matching, merging, and publication of **DQM** data.
* Performed match/merge and ran match rules to check the effectiveness of **DQM** process on data.
* Row profiling & column profiling by creating business rules on source tables using **Informatica Business Analyst.**
* Having experience in development, maintenance, support and enhancement in

**Informatica Power Center**, **Teradata**, Informatica **Data Quality** (IDQ), **Informatica**

**Big Data Management** (BDM), Oracle, Hadoop Hortonworks, Informatica Data Validation Option (DVO).

* Experience in implementing Data Quality rules on Hive Data sources using Informatica Big Data Edition 10.x
* Generating weekly score cards out of build profiles to monitor data health in **Informatica Business Analyst.**
* Development of UNIX shell scripts and for automation of ETL processes.
* Interacted with the Business Users to analyze the business requirements and transform the business requirements into the technical requirements.
* Follow the quality standards and guidelines for configuration management, testing, project planning and execution.
* Developed the test scripts using SQL for the unit testing.

|  |  |
| --- | --- |
| **Ameriprise Auto Home Insurance, WI**  **Sr Lead ETL Developer** | **Mar 2016 – Aug 2017** |
| **PNC Bank, Cleveland, OH**  **Sr Lead ETL Developer** | **Oct 2014 – Feb 2016** |
| **Sentry Insurance, Stevens Point, WI**  **IDQ Developer** | **Mar 2014 to Sep 2014** |
| **Nfrastructure, NY**  **Sr. Data Warehouse Consultant** | **Nov 2011 to Feb 2014** |