

## **Data Engineer**

### **Fremont, CA**

#### **Work Experience**

### **Data Engineer**

zaplabs - Emeryville, CA

January 2018 to Present

In the capacity of **Data Engineer**, my responsibilities include curation and stewardship of our crown jewels **data** • Designing and maintaining ETL pipelines around large volume **data** ingestion, clean-up, and processing using python, Scala, Airflow, spark and AWS technologies like S3, Athena, Redshift, Glue etc.

- Build cost effective, serverless ETL pipelines to transfer huge **data** from one **data** source into redshift **data** lake using combination of Python/Scala-AWS Glue- s3-Redshift • Build customized ETL pipeline using Airflow to transfer **Data** from Oracle into S3, RDS and Redshift • Use AWS Athena for extensive **data** analysis and interpret the results by visualization using AWS Quicksight • Extensive re-engineering using python and **data** cleansing using pandas to feed clean **data** for machine Learning Models to predict performance of an agent and attrition rate of agents within Realogy Group • Ensure quality of real estate **data** set by **data** manipulation and cleansing using pandas and insert the clean **data** set into Redshift • Connect Redshift/Athena/S3 datasets into Quicksight and build dashboards for visualization • Preparing functional or technical documentation for **data** warehouse system

### **Senior Data Analyst - Client**

Wells Fargo - San Leandro, CA

April 2017 to December 2017

Primarily support **data** management and business rules when transferring **data** between systems within the **data** governance group.

- Ability to understand and design **data** models and business Intelligence reports from **data** warehouses • Profile large **data** sets using complex SQL, Python libraries like numpy, pandas and translate it into rule mapping • Used statistical technique for building hypothesis testing to validate **data** and interpretation • Evaluates and designs **data** mapping logic to support governance and overall **data** quality management goals.
- Ability to write BRDs to transform complex **data** issues as exceptions • Work with our internal customers to understand, analyze and document business requirements.
- Support internal customers in resolving application issues, assessing user needs, providing training and recommending system changes

### **Computer Programmer**

3.International Computerware - Berkeley, CA

May 2016 to April 2017

Integrated Python with Oracle database and loaded **data** sets into AWS S3 • Apply various transformation logic by using Informatica to load **data** from legacy DB2 database to current Oracle 11g database • Use PL/SQL to convert **data** from legacy systems and load it onto Campus Solution website • Translate complex Business Requirement Document into technical mapping document for development purpose

### **Technology Lead - Client**

Infosys - Chennai, Tamil Nadu

March 2015 to March 2016

Applied various statistical methodologies using Python to study complex **data** sets • Analyzed large structured, unstructured and semi structured financial **data** sets using Python • Worked on Big **Data** Hadoop framework and developed and built hive tables using HiveQL • Worked on Hive optimizations, using partitioning and bucketing • Importing and exporting **data** from RDBMS into HDFS and vice-versa using sqoop • Developed backend application for clients Insurance portal using advanced PL/SQL concepts.

- Exclusively worked on UNIX environment and rewrote complex UNIX shell scripts as per the requirement.
- Wrote JIL scripts and scheduled jobs in Autosys • Used excel sheet, flat files, CSV files to generate Tableau adhoc reports and connected it to Tableau server to publish dashboard to a central location for portal integration.

### **Senior Associate Projects**

Cognizant Technology Solutions - Chennai, Tamil Nadu

December 2010 to February 2015

Analyze Potential customer data to target them for new product launch using Python • Format text files using Informatica to extract, transform, load data into HDFS, Hbase, Hive.

- Used advance features of PL/SQL for implementing complex business rules for clients auto and property insurance portal • Developed Unix Shell Scripts to automate backend jobs • Designed a three-tier database system for client's backend application to store Insurance data.

### **Senior Engineer**

Mindtree LTD - Bengaluru, Karnataka

June 2010 to November 2010

Involved in writing back end Packages extensively in PL/SQL.

- Maintained data integrity and security using Integrity Constraints and Triggers.
- Performance tuning of existing sql and pl sql codes to improve the response time.
- Created, debugged and modified Stored Procedures, Triggers, Tables, Views and User-Defined Functions.
- Resolve production defects using Service Now adhering to strict SLA

### **System Engineer**

Polaris Software Labs Limited - Chennai, Tamil Nadu

August 2007 to June 2010

Involved in writing back end Packages extensively in PL/SQL.

- Debugging existing PL/SQL code for better performance • Created, debugged and modified Stored Procedures, Triggers, Tables, Views and User-Defined Functions.
- Resolve production defects using case tracker tool adhering to strict SLA

### **Education**

Bachelor of Engineering in Mechanical Engineering

Anna University Chennai, Tamil Nadu

May 2006

### **Skills**

Data analysis (Less than 1 year), Database (Less than 1 year), Sql server (Less than 1 year), Sql server 2008 (Less than 1 year), Oracle (Less than 1 year), Sql (Less than 1 year), Hadoop(Less than 1 year), Hbase (Less than 1 year), Hive (Less than 1 year), Pig (Less than 1 year), Python (Less than 1 year), Matplotlib (Less than 1 year), Numpy (Less than 1 year), Pandas(Less than 1 year), Scripting (Less than 1 year), Cassandra (Less than 1 year), Oozie (Less than 1 year), Sqoop (Less than 1 year), Hbase (Less than 1 year), Data governance (Less than 1 year)

### **Links**

<https://www.linkedin.com/in/sriram-krishnamurthy-19523bb2>