

## CAREER OBJECTIVE:

To be associated with a progressive organization to utilize my skills and abilities in the industry that offers professional growth while being resourceful and Innovative.

## PROFESSIONAL SUMMARY:

- Dedicated associate software engineer with proven record in Designing, implementing, and maintain the data pipeline for analytical and reporting.
- Better Understanding of Hadoop Ecosystem and its major components
- Development and Implementation of various methods to load HIVE tables from HDFS and Local File System
- RDBMS Tables have been imported/exported using SQOOP.
- Developed Hive Queries to parse the raw data, populated external & manage tables and store the refined data in partitioned external tables.
- Worked in Hive Partitioning, Bucketing
- Experience in working with the Different file formats like TEXTFILE, ORC, Parquet and JSON and store it in HDFS/Hive tables using Spark.
- Hands on Experience on UNIX (Cloudera FTP, Web shell).
- Having good knowledge on AWS Cloud Technologies (S3, Lambda, Glue, Redshift).
- Having the best practice in data modeling, ETL process, and OLTP technologies.
- Worked on the ETL process for data extraction, transformation, and loading.
- Established the data quality processes, resulting in improved accuracy and reliability of organizational data.
- An exceptional performer with the ability to work in teams as well as good communication, analytical, relationship management and problem-solving skills.

## EDUCATIONAL QUALIFICATION:

- Completed B.E in Information Science & Engineering from Canara Engineering College Mangalore in the year 2022.

## TECHNICAL SKILLS:

- |                           |                                  |
|---------------------------|----------------------------------|
| • Language:               | SQL, Python, PySpark             |
| • Big Data Tool:          | Hive, HBase, HDFS, Spark, Sqoop  |
| • Databases:              | RDBMS, NoSQL                     |
| • Cloud Technologies:     | AWS (S3, lambda, Glue, Redshift) |
| • Reporting Tool:         | Tableau                          |
| • Data Governance:        | MDM, Data Quality                |
| • Version Control System: | GitHub                           |
| • Operating System:       | WINDOWS, UNIX                    |

## PROFESSIONAL EXPERIENCE:

- Working as an Associate Software Engineer in Maveric Systems Pvt Ltd in Bangalore from June 2022 to till date.

## PROJECT DETAILS:

### Project 1: Client Review Center

**Description:**

CRC is an efficient, User friendly, reporting engine that gives you access to robust library of Merrill Lynch approved exhibits – including information on performance, holdings, and asset allocation. The result is a colorful, easy to read professional report package that will help strengthen client relationship. The report generated using CRC provides the details about Cash, Equity, Fixed Income, Holdings, Mutual Funds, Alternative Investments etc. with Market value and percentage of accounts. The data will be persisted into the Hadoop file system, Hive are injected using Sqoop and Processing using Hive and Spark. Spark SQL load the aggregated data to RDBMS, Hive and Elastic Search for Analytics team and visualization in CRC UI

**Responsibly:**

- Data Ingestion and acquisition is done through Sqoop from RDBMS into hive table.
- Handling data sets and read data different source files like CSV, Parquet, Avro, Json, Sequence files and store it in HDFS/Hive tables.
- Importing historical transactions and user information into HDFS.
- Creating external and managed Hive tables with appropriate buckets and partitions.
- Data Enrichment, cleansing and common data aggregations through RDD transformations using Spark.
- Interactive analysis of Hive tables through various data frame operations using Spark SQL.
- Data parsing and processing by performing In-memory computation, data persistence, cache, repartition, performance tuning using Spark.
- Providing enriched data availability to frontend team.
- Storing the final processed data into RDBMS, Hive table and HDFS for making the final data available for visualization in CRC UI which the Analytics/Reporting teams can make use of it.
- Collaborate with the cross-function teams to understand and address data requirements.
- Involved in story-driven agile development methodology and actively participated in daily scrum meetings.

---

**PERSONAL STRENGTH:**

- Quick learner and adaptable to new environment.
- Smart working, Determination, Innovativeness.
- Good Interpersonal and Communication skills.

---

**DECLARATION:**

I hereby declare that the above-mentioned information is correct up to my knowledge and I bear the responsibility for the correctness of the above-mentioned particulars.

Koushik N S

Place: Bangalore

Date: 5/01/2024