**SALMAN SAYYAD**

**+91-7517-564-738|python831547@gmail.com|**[**linkedin.com/in/salman-sayyad-88b016a**](https://www.linkedin.com/in/salman-sayyad-0b88b016a)

Big Data Engineer with 3+ year of experience leveraging ETL, Big Data Transformations, Aggregations, Data warehousing and reporting.

**Professional Experience**

**Ample SofTech -** Data Engineer.

**Client -** Swiggy : Apr 2018 - Present

**Project -** E-commerce

Swiggy is India's largest and highest-valued online food ordering and delivery platform. Logs generated by millions of users through Swiggy website and mobile application saved in AWS S3 bucket through Kafka server. Project involves ETL, transformation, aggregation, data warehousing and reporting using batch and stream processing.

**Technology**

**Spark 2.4.x, Hadoop 2.7.x, AWS S3, EMR, Scala 2.12.x, Kafka, Sqoop, SQL, Hive, HBase, Cassandra, Python pandas, numpy.**

* Batch and stream processing of data through kafka and spark.
* Worked with different file formats like json,avro,csv,orc,parquet etc.
* worked on hive using performance techniques like partitioning and bucketing.
* Performed ETL, Data wrangling, data analysis using Scala, Spark.
* Worked on ksql to perform aggribations on streaming data.
* Reading and writing data to and from kafka topics.
* Very good undersanding of kafka architecture and it’s working
* Developed Spark code using Scala and Spark-SQL for faster processing of data.
* Experienced in handling large datasets using Partitions, Spark in Memory capabilities, Broadcasts in Spark, Effective & efficient Joins, Transformations and other during ingestion process itself.
* Developed Scala scripts, UDF's using both Data frames/SQL and RDD/MapReduce in Spark 2.4.x for Data Aggregation, queries and writing data back into RDBMS through Sqoop.
* Involved in complete end to end code deployment process in Production.
* Worked on back-end using Scala 2.12.0 and Spark 2.4.0 to perform several aggregation logics.
* Experience in working with NoSQL databases like Hive 2.1.1 ,HBase and Cassandra.
* Very Good understanding and Working Knowledge of Object Oriented Programming (OOPS).
* Excellent analytical, Interpersonal and Communication skills, fast learner, hardworking and good team player
* Very good understanding of Machine Learning and Deep learning algorithms like Linear regression, Logistic regression, SVM, KNN, Decision Tree, Random Forest, Backword propagation etc.

**Infosys BPM -** Process Executive /Sales Reporting Executive

**Client -** Chobani : May 2016 – Mar 2018

**Project -** E-commerce

Chobani is an American food company specializing in strained yogurt. Project involves analysis and reporting of daily sells transaction data using SQL Server and Power BI.

**Technology**

**Power BI, Sql, Microsoft SQL Server**

* Imported data from SQL Server DB to Power BI to generate reports.
* Created Dax Queries to generated computed columns in Power BI.
* Generated computed tables in Power BI by using DAX.
* Involved in creating new stored procedures and optimizing existing queries and stored procedures.

**Academic Qualification:**

**Dr. Babasaheb Ambedkar Marathwada University, Aurangabad. *B.Sc. Computer Science* Aug-2014**