

# Ranjithkumar B

Technology Specialist- Bigdata,Aws  
Data Engineer

## CONTACT

☎ 8754565484

✉ [ranjith.baskaran05@gmail.com](mailto:ranjith.baskaran05@gmail.com)

📍 1/419 pillaiyar kovil street,  
kelambakkam,  
Chennai – 603103

## SKILLS

Python

Spark - Core, SQL

Hadoop - HDFS, Hive

Sqoop

Apache kafka, Airflow

SQL, Oracle

Jira, winscp, putty

AWS - EC2, S3, Redshift,

EMR, Glue, Athena,Lambda

Autosys

Shell script

## PERSONAL STRENGTHS

Critical and Analytical Thinker

Quick Learner

Problem Solver

Optimist

Go-Getter

## EDUCATION

BE – Electronics Engineering

Dhanalakshmi Srinivasan College  
of Engineering, Chennai

2013 – 2017

HSC- St Joseph Higher Secondary  
School, Chennai

2011 – 2013

## Summary

- Passionate and Innovation oriented **Data Engineer** with 3+ years of hands-on experience.
- Currently seeking a challenging position in a reputed technical organization where I can enhance my skills and strength in conjunction with the institution's goal and achieve excellence through experience.

## Roles and Responsibilities

### **Associate Data Engineer**

*Publicis Sapien, Bangalore / June 2022 – Present*

- Gathered the business requirements from the Business Partners and Subject Matter Experts
- Developed **Spark SQL** code in Spark as per client requirements.
- Experience in managing and reviewing **Hadoop log files**.
- Experience writing unit tests for Business requirements.
- Strong analytical, troubleshooting, and problem-solving skills
- Developed spark data frame to loading the **hive** aggregate table to **Aws redshift**
- Developed various **shell scripts** for running the spark submit to triggering the jar file
- Created jill file and scheduled the jobs in **Autosys** and monitored the result

**Tools and Technologies:** Big Data Ecosystems, Spark/PySpark, Hive, Unix, AWS REDSHIFT,Autosys

### **Associate**

*Cognizant Solutions, Chennai / Dec 2018 – June 2022*

- Worked closely with business customers for Requirement gatherings
- Bulk importing of data from various data sources like **Aws S3 and Oracle** into **HDFS** using **Spark JDBC** and **Boto API**
- Using HDFS for persisting the data for **primary staging**
- Creating **Hive** tables with appropriate **partitions** and mapping with ingested data.
- Developed **Spark SQL** to process the data and performing transformation and loading into the **hive table**
- Good Knowledge in writing different kinds of tests like **Unit test cases** In SQL and Hive
- Developed **Autosys jill file** for scheduling to run the **spark submit** command
- Coordinate with onsite business analyst and application developers, attend client meetings.

**Tools and Technologies:** Big Data Ecosystems, Spark/PySpark, Hive, Oracle, Unix, AWS S3, Autosys