

Geo Thomas

Desirous of a Data Engineer Position to utilize my technical and management skills in a challenging workspace that welcomes innovative ideas and offers opportunities to enhance my knowledge in emerging trends in the IT sector.

Mob : +91 8547130930

Mail : geothomas1010@gmail.com

[Linkedin.com/in/geo-thomas](https://www.linkedin.com/in/geo-thomas)

EXPERIENCE

Tata Consultancy Services,Kochi — Software Engineer

March 2018 - PRESENT

- Currently working as a Scala/JAVA/Spark Developer.
- Hands on experience in designing and implementing applications using Java,scala,airflow, hadoop components and aws resources.
- Have knowledge on Hadoop and Big Data concepts.
- Expertise in Hive and mysql query generation and analysis.
- Hands-on experience in CI-CD build and deployment process.
- Hands-on experience in Agile Methodology of Software Development

SKILLS

Spark

Scala

SQL

Java

Airflow

AWS

Hive

Hadoop

Shell Scripting

Maven

Gradle

PROJECTS

Media Data Lake - May 2018 -April 2020

Data lake for all the data coming from TV audience measurement. Data from different sources(Netezza, Oracle,Redshift,S3) are ingested into the data lake(S3 bucket) with metadata in the apache hive metastore.

Role & Responsibilities:

- Development of components to ingest data from various sources and sync metadata from external hive and athena metastore to data lake metastore.
- Development of a api based utility to copy subset of production data to non prod environment
- Created and scheduled jobs using Airflow.
- Aws account migration and aws emr version migration
- Data quality validation

TOOLS USED

Eclipse | IntelliJ | Jenkins | Jira

Putty | Git| Winscp

DBVisualizer | Notepad++

Technology: Hadoop Ecosystem (Hive, Spark, Oozie), Java, SQL,AWS,Airflow,Shell Script

Total Content Rating — *May 2020-Present*

Total content rating is the measurement of TV audience data at different levels. Commercial data is calculated and removed from the data.

Role & Responsibilities:

- Rewriting on-premise SQL job runs odi environment to cloud environment using scala,spark,airflow,aws resources.
- Development of Tv Commercial data calculation
- Development of Out of home viewing data reporting
- Performance tuning.
- Data quality validation

Technology: Spark, Scala AWS- (S3,EC2,SNS,SQS,Redshift,Secret Manger),airflow,SQL

AWARDS & ACHIEVEMENTS

Best Team Award - 2018

TCS Digital cadre Associate

EDUCATION

MES College of Engineering, Kuttippuram — B.Tech in Information Technology.

August 2013 - May 2017

Completed B.Tech in Information Technology with a **64%**

MTHSS,Chungathara — Plus Two

June 2011 - March 2013

Completed higher secondary education (Plus Two) with **81% marks.**

GHSS, Moothedath— SSLC

June 2010- March 2011

Completed SSLC with **81% marks.**