

# BABJI LOGANDA

✉ [babjimanikandan@gmail.com](mailto:babjimanikandan@gmail.com)

☎ +1 (438) 860-0243

💼 [linkedin.com/in/babji-loganda](https://www.linkedin.com/in/babji-loganda)

📍 Toronto, Canada

## TECHNICAL SKILLS:

- ✦ **Big Data Framework/ Tools** : Spark, Hadoop, Hive, Kafka, HDFS, Delta, Flink, SQOOP
- ✦ **Languages** : Java, Scala, Python, SQL
- ✦ **Cloud Data Platforms** : AWS, DataBricks, Cloudera, GCP
- ✦ **Databases** : Postgres, MySQL, Oracle, MongoDB
- ✦ **Data Quality** : Drools, Great Expectations, AWS Deequ
- ✦ **Build, Version Control** : SBT, Maven, Jenkins, GIT
- ✦ **Scheduler & Monitoring** : Airflow, DALM

## EDUCATION:

- **Bachelor's in Electrical and Electronics Engineering** | Anna University **Aug 2010 - Apr 2014**
- **PG Diploma in Business Administration** | Annamalai University **July 2020 - Jun 2021**

## PROFESSIONAL EXPERIENCE:

**Locala | Lead Data Engineer | Toronto, Canada** **1 yr 7 mos | Apr 2022 - Present**

- Built large-scale data ingestion and denormalization pipeline of advert data extracted from APIs and file storages using spark batches to compute user clicks, displays, footfalls, and cross-visitations.
- Additionally, I built near real-time data reconciliation with spark structured streaming to assist brands in targeting their audiences and increasing foot traffic to their stores/platforms.
- Ensured Data Quality on each of the data pipeline using Great Expectations and AWS Deequ.

**Tech Stack:-** Spark, Scala, Java, Python, Hive, Delta, Postgres, DataBricks, Airflow, Great Expectations, Deequ

**PayPal | Data Engineer 2 | Chennai, India** **2 yrs 10 mos | May 2019 - Mar 2022**

- As a Big Data Engineer my primary responsibility was to create an authoritative data store by sourcing data from transactional databases of payment transaction data, profile and clearing loads.
- I ensured data availability for downstream finance applications to build sub-ledgers for transactional audits, settlements, and reconciliation on inbound and outbound transfers. The ledgering framework was built using spark ETL pipeline which caters the purpose of book balance on hourly and daily basis.
- Improved data-quality by designing and building a Data Quality Framework using Drools for validating large scale enterprise data both real-time and in batch.

**Tech Stack:-** Spark, Java, Hadoop, Hive, Kafka, Oracle, Cloudera, Airflow, GCP, Drools

**Mindtree | Senior Engineer | Chennai, India** **4 yrs 9 mos | Oct 2014 - May 2019**

- Built a data ingestion pipeline that generated analytical insights on booking and reservation data. This pipeline was designed to overcome the limitations of traditional mainframe systems, which were often slow and inefficient in data processing.
- Implemented a wide network data management system to keep track of Products stocking, sales reports and product delivery and new launches based on sales summary.
- Worked on data comparison framework which helped the client to build a Location proximity comparison engine for the user navigation and mapping.

**Tech Stack:-** Scala, Java, Spark, AWS, Hadoop, Kafka, Hive, Postgres, MySQL

## PROJECTS & CERTIFICATION:

**Data Validator:** Built data quality framework for Spark streaming and batch application which can validate datasets inline(in memory) and offline for tables and files and also used to validate entities in java applications. The framework has option to develop rules based on business needs.

- <https://github.com/babjiloganda/DataValidator>

HDP Certified Spark Developer - <https://bcert.me/sltxeszw>

**May 2017**

## Awards & Recognition:

**Mindtree** Pillar & outstanding performer Award for 2015, 2016, 2017 & **PayPal** A team award