

Nikam Amit Tukaram

+91-9029896958

nikama254@gmail.com

Big Data Engineer

Infosys, Pune [From Sep, 2019 to till date]

PROFILE

“**Big Data Engineer**” with 3+ years of experience in Big data/Hadoop technologies. Well-acquainted with Data warehouses, Data pipelines, Data structures, Data managements and processing systems. Strong desire to move forward, face new challenges and expand my skill set.

PROFESSIONAL SUMMARY

- Strong experience in **Hadoop** and **Spark ecosystems** including **Spark core, Spark-sql, HDFS, Hive, Sqoop**.
- Strong background with file distribution system in a big data arena.
- Good knowledge of **Spark architecture** and its integrations like RDD, Dataframe and Dataset.
- Having good knowledge on programming language as **Python** and **Scala**.
- Proficient in performing **SQL queries**.
- Easily adopt, explore, learn and understand newer business domains and technology based on the requirements.
- Develop best practices for developing and deploying **Hadoop applications** and assist the team to manage compliance to the standards.
- Good knowledge of **python programming** and **object oriented programming** concepts.
- Developed jobs for ingesting tables in Hive and from Hive to Big Query.
- Manage and review **Hadoop log files**.
- Perform tasks such as **writing scripts, calling API's, write SQL query** etc.
- Hands-on Experience in mainframes and Hadoop eco-system. Good exposures in Requirement Analysis, Customization and user acceptance.

TECHNICAL SKILLS

Type/Name	Details
Operating Systems	Windows Family, Linux
Database	MySQL
Hadoop system	HDFS, Map-Reduce, Hive, Sqoop, Spark-core, Spark-sql, Hbase.
Programming Language	Python, Scala, Pyspark, SQL

EDUCATIONAL QUALIFICATION

Bachelor of Engineering (May 2019) from Mumbai University

PROJECT DETAILS

Project 1:

- **Project name:** FMO, Netherland
- **Domain:** Banking
- **Tools:** Hadoop, Sqoop, SQL, Hive, Hbase

Responsibilities:

- Used Sqoop to pull data & created Hive tables on data in raw layer.
- Implemented delta load functionality to maintain new, updated & history records in table.
- Final table inserted into enriched layer in hive.
- Fully pipeline handled using shell script.
- Successfully completed POC on HBase.
- Involved in extension of pipeline from Hive to HBase for dashboard/visualization.
- Extracted the data from MySQL into HDFS using the Sqoop.
- Created Hive tables, dynamic partitions, buckets for sampling and working on them using Hive QL.
- These data get transferred into 3 different Hive tables based on each logic.

Project 2:

- **Project name:** Elverys, Ireland
- **Domain:** Retail
- **Tools:** Spark-sql, Spark-core, Pyspark, Hbase, Hive

Responsibilities:

- Sources for this project were MYSQL database, which gets combined with file sources
- Which need to combine by using ETL operations.
- Completely involved in the requirement analysis Phase.
- Involved in Pyspark programming for preprocessing and cleaning of data for next stage of analytics.
- This tables were then transferred into raw to processed to curated forms via Python/Spark SQL scripts, which later loaded into Hive.
- Performed bulk load from Hive to Hbase.

PERSONAL DETAILS

Date of Birth : 16th July 1995

Marital Status : Single

Languages : English, Hindi & Marathi

Address : Chinchwad, Pune, Maharashtra.

DECLARATION

I do here by declare that the above mentioned information is true and complete to the Best of my knowledge and belief.

Place: Pune