

# Narayan Mhaske

SOFTWARE ENGINEER

## ABOUT ME

I have total **6 years** of experience in IT Industry. Handle Big Data using **Google Cloud Platform, Apache Spark, SQL, Python**. Good team player as well as able to prioritize own workload and work under own initiative.

## PROFESSIONAL SUMMARY

- Experience in working on GCP services - **Big Query, Google Cloud Storage, Google Cloud Composer, Dataproc**.
- Hands on experience on **Apache Pyspark**.
- Work with **Spark data frame** operations and aggregate functions, window functions, analytical functions.
- Experienced in implementing **Spark RDD transformations and actions**.
- Strong experience in implementing Spark code to perform **ETL**.
- Involved in writing data cleansing and transformations using Apache Spark operations.
- Experience with **RDD, Spark Data frames, Spark SQL**.
- Knowledge about HIVE concepts as per the requirement.
- Experience in writing complex SQL queries **using joins, functions, sub queries, correlated sub queries**.
- Dealing with various file formats like CSV, Parquet, ORC, Avro, JSON.
- Experience of creating data pipelines with Cloud Dataproc.
- Experience of creating **DAG** with the help of cloud composer.
- Working experience on **GCP partitioning and Clustering**.
- Good knowledge of Software Development Life Cycle and Extensive experience working in an Agile development environment.
- Extensive knowledge and working experience on CI/CD using tools like GIT.
- Familiarity with delivery methodologies like Agile
- Running analytics queries.
- Good Experience of core python and concepts like List Comprehensions, Regular Expressions, decorators, generators, OOPs, Functions.
- Good documentation skills and technical Skills.
- Team player with interpersonal and good communication skills.

 +91-8767630224

 narayanmhaske05@gmail.com

 Pune

## RELEVANT SKILLS

### Operating Systems

- Windows
- Linux

### Big Data (GCP Tools & Services)

- BigQuery
- Dataproc
- Compute Engine
- Google Cloud Storage
- SQL
- Cloud Composer(Airflow)
- Google Cloud SDK Shell
- PYSPARK
- Data-Processing

### Programming Languages

- SQL
- PYTHON

### Version Control Sys.

- GIT

## EXPERIENCE SUMMARY

### Project\_3

---

**Client:** Delta Air Lines

**Role:** Software Engineer

**Description:-** Delta Air Lines is one of the major airlines of the United States and a legacy carrier. It is ranked second among the world's largest airlines by number of passengers carried, passenger-miles flown, and fleet size. It is ranked first by revenue for commercially owned airline companies, Delta is headquartered in Atlanta, Georgia.

- Created python code for various google cloud API clients like BigQuery, Configuration, Dataproc and Created Python utilities which will be used to connect source and target systems.
- Creation of various scripts to automate the activities like data loading.
- Works on existing PySpark code and tried to make them more generic.
- Created Python dataflow templates to process large volume of data.
- Performing schema validations, special character validations and data cleansing using Spark.
- Develop SQL Queries for fetching the data from database.
- Coordinating with the team to meet the company requirements.

### Project\_2

---

**Client:** Caltex Oil

**Role:** Software Engineer

**Description:** Caltex is a petroleum brand name of Chevron Corporation used in the Asia-Pacific region, the Middle East, and Southern Africa. It is also the brand name of non-Chevron petroleum companies in some countries (such as New Zealand, and previously Australia and South Africa) under a trademark licensing agreement with Chevron.

- Performing pre-processing and cleaning on data using Spark.
- Involved in GCP architecture implementation for a complete cloud based big data solution using Bigquery, GCS, Dataproc and Airflow.
- Responsible for pull, push and committing the code using GITHUB.
- Debugging and fixing the issues by sampling the data.
- Sorting the extracted data in Oracle database for further querying and analysis.
- Responsible for debugging and troubleshooting during data processing.
- Involvement in analysis of the project requirement.

### IDEs

- Jupiter Notebook
- Pycharm

### RDBMS

- Oracle

### EDUCATION

BSC (Computer science) in Oct 2017 from **Pune University, Maharashtra**

### WORK EXPERIENCE

Worked with Tech Mahindra Ltd. from July-2016 to July 2018

Worked with WNS global services Ltd. from May-2019 to July-2022

Working with ICONIC E-TECH from Aug-2022 to till date

## Project\_1

---

**Client:** MAPFRE Insurance

**Role:** Software Engineer

**Description:** MAPFRE Insurance is headquartered in Webster, Massachusetts, United States. It was founded in 1972 as the Commerce Group it was acquired in 2008 by MAPFRE S.A. . MAPFRE Insurance writes property and casualty insurance in 19 states across the United States through a network of more than 4,200 independent agents and brokers.

- Develop SQL Queries for fetching the data from database.
- Responsible for pull, push and committing the code using GITHUB.
- Debugging and fixing the issues by sampling the data.
- Sorting the extracted data in Oracle database for querying and analysis.
- Responsible for debugging and troubleshooting during data processing.
- Writing complex SQL queries using joins, functions, sub queries and correlated sub queries fetching the data from database.