

# Shubham Dubey

[dubeys7151@gmail.com](mailto:dubeys7151@gmail.com) | [Personal Website](#) | [LinkedIn](#) | [GitHub](#)

## Education

SRM Institute Of Science and Technology B.Tech, Computer Science and Engineering (2020)

## Skills and Tools

Programming Language : Python, Java  
Data Engineering : Apache Hadoop, spark, Hive, SQL (mysql)  
Devops: Git, Jenkins

## Certificate

Big Data Hadoop and Spark Developer, SimpliLearn | [View](#)  
Python Introduction, Youtube  
SQL, Youtube

## Language

English  
Hindi

## Experience

Senior System Engineer, INFOSYS LIMITED — ( 4/2022 - Present )

- Worked in the Quality Engineering team for client Apple Inc.
- Contributed in testing and support of Corp applications.
- Worked on testing and successful deployment of FIDO features.
- Analyze product requirements and come up with a test plan.
- Implemented, tested and automated scenarios with over 80% coverage.  
Tech Stack : Java, TestNg, git, Jenkins, SQL (Oracle DB)

System Engineer, INFOSYS LIMITED — (01/2021 - 03/2022)

- Product owner for corp features like group, person service, SCIM.
- Regression Analysis was part of the day to day task.
- Worked efficiently under an agile environment with Devops tools.  
Tech Stack : Java, Custom Testing Framework, git, Jenkins

System Engineer Trainee, INFOSYS LIMITED — (01/2021 - 03/2022)

- Trained as Full Stack Web Developer specialized in Mean Stack.
- Lead Team of 4 to successfully complete an end to end project.
- Scored 83% and was classified as High Performer.

Database Management Intern, Uflex LIMITED — (12/2018 - 01/2019)

- Supported in Data entry and analysis.
- Report generated as part of excel sheet.

## Project

Movie Data Review using Apache Hive | [View](#)

- Project focuses on loading data into Apache hive and analyzing the same.  
Source: <https://simplilearn.com>
- The data is collected from IMDB, with 49000+ records. The records are analyzed by querying using Hive query Language.
- Used Hive UDF to convert data from one form to other to facilitate better analysis  
Tech Stack: HDFS, HIVE, Python, HQL, matplotlib

Log Analysis using Spark | [View](#)

- Project Focuses on analysis of log using spark Dataframe on data-bricks  
Source : EDGAR log file data set
- The data have over 1 lakhs records, loading the data to data-bricks and using the platform to analyze the data
- Used spark UDF to convert data to different types.  
Tech Stack : Python, Apache Spark, data bricks, spark sql, spark dataframe

