

PROFILE

Seeking a challenging position in Big data technologies to utilize my skills and abilities while being resourceful, innovative and flexible.

Overall 9+ years of experience in IT, and 4+ years of experience in Big data Technologies. Experience working in Hadoop, Spark and AWS (cloud platform).

CONTACT

PHONE:

+27 67 218 6503

Linked-In Profile:

www.linkedin.com/in/shashi-dataengineer-jnb

EMAIL:

Shashireddy40@gmail.com

PORTFOLIO

<https://shashireddy40.github.io/mypage/>

CERTIFICATIONS

ISTQB-CTFL

Big Data Engineering with Hadoop and Spark from E&ICT, IIT Roorkee

PERSONAL PROFILE

Full Name:

SHASHIDHAR REDDY

Date of Birth:

10th May 1990

Languages Known:

English, Telugu and Hindi

ACHIEVEMENTS

Employee of the Quarter FY17

HOBBIES

Chess

Cricket

Traveling

SHASHIDHAR REDDY

Data Engineer (Hadoop,Spark and Python)

Johannesburg,SA

EDUCATION

Jawaharlal Nehru Technological University,Hyderabad,India.

Bachelor of Technology in Aeronautical Engineering.

WORK EXPERIENCE

Eblocks Software (Client OLDMUTUAL SA)

Data Engineer

(Oct 2019-May 2020)

Sonata Software, India.

Big Data Engineer (4 Years)

January 2015–January 2019

Hadoop and Spark (Big data Specialist)

IrusInfotech Pvt Ltd, India

QA Analyst (3+ Years)

June 2011–January 2015

ETL,SQL, Manual Testing

SKILLS

Operating System:

Linux,Windows

Languages:

- Python
- Scala
- Shell Scripting

RDBMS:

- Oracle SQL Server,MySQL

Big data Technologies:

- Hadoop,Hive,Sqoop
- Spark
- Cloudera/Hortonworks Distribution

Utilities/Tools:

- IntelliJ,Docker, API Tools,Bitbucket.

Cloud Platforms:

- AWS(S3,EMR,EC2,RDS)

MS-BI

- SSIS(ETL)

Data Visualization Tools

- Tableau,Matplotlib

PROFESSIONAL EXPERIENCE

PROJECT #1:

Title: CBR-R (CONDUCT OF BUSINESS RETURN REPORT)
Client: OLDMUTUAL SA.
Location: Sandton, Johannesburg.
Duration: OCT 2019 to Present
Tools Used: Hadoop, SQL, Hive, Apache Sqoop, Hortonworks
Team Size: 5
Role: Data Engineer.

Description:

The Conduct of Business Return must be submitted to the Financial Services Board which is the Regulatory Authority for the CBR. The CBR is applicable to all Insurers carrying on either a Long-Term Insurance Business, or a Short-Term Insurance Business, or both. It calls for metrics on a number of aspects, dealing with the conduct of the insurer as it relates to policyholders. Policy numbers, new business, cancellations, surrenders, lapses, benefits, claims, complaints and commissions are the main focus areas.

Responsibilities:

- Design, create, build & maintain data pipelines
- Handling, transforming & managing Data using Big Data Frameworks.
- Building complete infrastructure to ingest, transform & store data for further analysis & business requirement.
- Taking care of the complete ETL (Extract, Transform & Load) process. Like Data ingestion, Data Transformation Using Apache Sqoop, Spark and Hive.
- Generating Reports for FSB(Financial Board of South Africa) Using Hadoop ecosystems.

PROJECT #2:

Title : TUI Travel Group
Brands : Thomson Airways, Amadeus System, British Airways.
Client : TUI UK, British Airways.
Location : Hyderabad, India.
Duration : Nov 2015 to Dec 2018
Tools Used : Hadoop, Hive, Spark, Sqoop, Cloudera, Tableau
Cloud Technologies : AWS S3, EC2
Team Size : 10
Role : Big Data Engineer.

Description:

TUI Group is a multinational travel and tourism company. It is the largest leisure, travel and tourism company in the world, and owns travel agencies, hotels, airlines, cruise ships and retail stores. TUI Travel Group have several UK-based tour operators Thomson Holidays and First Choice, by using this brands Passengers able to Book Flight tickets, Hotels and Holiday Packages with Cheapest Prices to Major European Countries

Responsibilities:

- Designing and implementing semi-structured data analytic platform leveraging Hadoop.
- Extensively used Hive/HQL or Hive queries to query or search for a particular string in Hive tables in HDFS.
- Worked on analyzing Hadoop stack and different big data analytic tools including Pig, Hive, and Sqoop.
- Implemented Spark using Scala and Spark SQL for faster processing of data.
- Used Spark for interactive queries, processing of streaming data and integration with popular NoSQL database for huge volume of data.
- knowledge in NoSQL databases like MongoDB
- Exported the analyzed data to the relational databases using Tableau for visualization and to generate reports for the BI team.
- Developed recommended system with using Machine learning techniques such as scikit-learn and python libraries.

PROJECT #3:

Title	: HR Berry– workforce and Payroll management Application
Client	: Ascent Consulting Services Private Limited, India
Location	: Hyderabad, India
Duration	: Feb 2015 to Oct 2015
Tools Used	: Eclipse IDE, Python, Hadoop, MySQL, Sqoop,Docker
Team Size	: 7
Role	: Big data Specialist.

Description:

HR Berry is an on-premises human capital management (HCM) solution suitable for large organizations with more than one thousand employees. AscentHR HRberry offers integrated tools to manage recruiting, talent management, time management, and employee onboarding. Employee and Manager Self Service Global Payroll Processing application PowerHR & PowerDox and make this a unique proposition.

Responsibilities:

- Acquire data from primary or secondary data sources and maintain databases/data systems
- Operate and optimize pre-defined tools, applications and data bases/data management systems
- Process, manage, extract and cleanse data to apply Data Analytics in a meaningful way
- Create reports and communicate results to various internal and/or external stakeholders
- Worked on Data Frames and Complex data types.
- Exploring and analyzing data, performing Grouping, Aggregation and working with joins.
- Querying data using SQL and Data Frame operations
- Developed a frame work to handle loading and transform large sets of unstructured data from UNIX system to HIVE tables.
- Involved in Data Ingestion to HDFS from various data sources.
- Extensively used Apache Sqoop for efficiently transferring bulk data between Apache Hadoop and relational databases
- Worked extensively with importing metadata into Hive and migrated existing tables and applications to work on Hive and AWS cloud.

PROJECT #4:

Title	: eCRS (electronic Credit Rating System)
Client	: National Australia Bank, Australia.
Location	: Bangalore, India.
Duration	: Jan 2013 to Dec 2014.
Tools used	: Eclipse IDE, SQL Server, SoapUI, Postman
Team Size	: 5
Role	: Data Analyst

Description:

eCRS is an online Credit Rating System that helps the bankers/employees of NAB to take a decision about issuing loans to their respective customers (organizations).

There are several external factors involved in this project in order to drive the banker to take a proper decision against issuing the loans to their customers. It's a web application with different roles aligned with different privileges to access the system. The data from this system would be pushed down to another project called GDW. So the data extraction is a crucial process in this project.

Responsibilities:

- Developed SQL queries with various data sources like JSON, Parquet.
- Exploring and analyzing data, performing Grouping, Aggregation and working with joins.
- Querying data using SQL operations.
- Responsible for test scripts generation, testing and validation.
- Have a good understanding of Software Development Life Cycle (SDLC), involved in various phases like Requirement gathering, Analysis/Design, Development and Testing. Also knowledge in Test Driven Development (TDD) process.

PROJECT #5:

Title	: Grocer-kart
Client	: Idea Logistics, Hyderabad.
Location	: Hyderabad, India
Duration	: June 2011 to November 2012
Tools Used	: Quality Center, QTEST, SQL Server 2008.
Team Size	: 5
Role	: Senior Test Analyst (Manual, Database Tester)

Description:

Grocer kart is an online shopping portal they do save on the infrastructure costs that we would otherwise have to incur if we had a physical store. Since at Grocer kart, the customer always comes first, we have decided to pass on our savings to you in terms of discounts on products. Customer will realize that most products are priced lesser than the MRP.

Responsibilities:

- Have a good understanding of Software Development Life Cycle (SDLC), involved in various phases like Requirement gathering, Analysis/Design, Development and Testing.
- Responsible for test scripts generation, testing and validation.
- Analyzing Functional Requirements, Gathering and managing the Test Data
- Effective test case writing and execution resulting in improved coverage
- Providing input/status updates regarding test progress/results
- Assist in the assessment and planning of test efforts required for new functions/features under development
- Designing Test cases using HP Quality Center and Executing Manually.
- Defect tracking, analyzing and bug reporting using QC.
- Defect and trend tracking, reporting, and analysis