



RISHI RAJ RAJPUT

"Expertise in applying Big Data Technologies like PySpark, Hadoop, Hive Platform, Sqoop and Cloud Technologies"

"High interest and experience in closely following technological developments in the marketplace relating to big data and customer intelligence"

Mobile: +91 9039151998

Email: rishirajrajput21@gmail.com

LinkedIn: www.linkedin.com/in/rishi03 **Git:** <https://github.com/RishiRajRajput>

PROFILE SUMMARY

- ❖ A competent professional with nearly 5+ years of experience in **Big Data & Hadoop Analytics Project Management, Operations Management, Process Management, SLA & Transition Management**
- ❖ Experienced in Hospital and Retail industries using tool like **Hadoop, PySpark, Python, Hive, Sqoop and Airflow**
- ❖ **Capabilities in interpreting data and providing ongoing reports**, implementing big data and analytics architecture and design including gathering requirement for customer elements to be incorporated for analytical models across different data sources - structured/semi-structured
- ❖ **Capable of managing business analysis functions** right from stakeholder interaction, requirement gathering, documentation, sharing requirement with the technical team, developing solutions and obtaining stakeholder sign-offs
- ❖ Hands-on experience in using tools (Advanced Excel & PowerPoint, Big Data Analysis both Structured & Semi-structured Data); skilled in project lifecycle management, reporting and supporting seniors in project status
- ❖ Gained overall experience in **extensive excel, best practice documentation, data validation and classifying the data into various categories**
- ❖ Experience in managing complete **Data Warehouse Development Life Cycle including Enterprise Architecture, Data Analysis, Database Design, Database Performance Tuning & Optimization, Data Warehouse Optimization**

CORE COMPETENCIES

- | | | |
|---|--|--------------------------------------|
| ❖ Big Data Analytics | ❖ Data Analysis & Reporting | ❖ Data Profiling, Data Cleansing |
| ❖ Data Management (Hadoop & Cloud Technologies) | ❖ Structured/Semi-structured Data Analysis | ❖ Stakeholder & IT Team Coordination |
| ❖ Requirement Gathering & Analysis | ❖ Hadoop Development & Implementation | ❖ Project Planning & Execution |
| ❖ Project Management | ❖ Agile Methodology | ❖ Team Building & Leadership |

TECHNICAL SKILLS

- | | | |
|----------------------|------------------|--|
| ❖ Python | ❖ SQL Server | ❖ Big Data Hadoop |
| ❖ Spark, Hive, Sqoop | ❖ Airflow, Kafka | ❖ AWS Clouds (S3, EMR, EC2, IAM, Databricks) |

WORK EXPERIENCE

Since Apr'22 with Affine Analytics, Karnataka as Senior Associate Data Engineer

Working on tools such as – Sqoop, Hive, PySpark, Kafka, Airflow, AWS Clouds (S3, MapR, IAM)

Key Result Areas (Roles and Responsibilities):

- Developing and maintaining scalable data pipelines and building out new API integrations to support continuing increases in data volume up to 12 GB and with high complexity
- Collaborating with analytics and business teams to improve data models that feed business intelligence tools, increasing data accessibility, and fostering data-driven decision making across the organization
- Performing data analysis required to troubleshoot data related issues and assisting in the resolution of data issues
- Designing data integrations and data quality framework
- Administering a wide range of Big Data and Cloud technology and tools to support and accelerate activities such as Hadoop (MapR 3.0), PySpark
- Planning and executing Agile Methodology and DevOps processes, and timely identifying and escalating observations, risks, and issues

Big Data Analytics

- Ensuring the successful delivery of emerging big data solutions and defining strategy, principles, vision, and standards; providing advisory services around the data frameworks, governance, and quality
- Simulating, designing, developing & deploying computationally complex and practical data; building and delivering comprehensive data strategy roadmaps

- Forming alliances with department managers on projects, operational decisions, scheduling requirements/conflicts, and vendor contract clarification; leading the development of enterprise technology standards, governance processes, and performance metrics to ensure company delivers value

Feb'22 to Apr'22 with Futureense Technologies Private Limited Karnataka as Data Engineer

Skills – Sqoop, Hive, PySpark, Kafka, Airflow, AWS Clouds (S3, EMR, IAM, EC2)
Roles and Responsibilities – Gather data from multiple sources and create a data pipeline with spark version 3.0

Apr'20 to Dec'21 with BSP Technologies, Gujarat as Senior Software Developer

Skills - Big data tools PySpark, Hive, Sqoop, SQL Server, PL/SQL, and AWS Clouds (S3, EC2)
Roles and Responsibilities – Gather data from multiple sources and create a backend data pipeline with spark version 3.0

May'19 to March 20 with Cognic System PVT. LTD., MP as System Analyst

Skills – Big data tools PySpark, Hive, Sqoop, SSIS, PL/SQL, and AWS EC2
Roles and Responsibilities – Gather data from multiple sources and create a backend data pipeline with spark version 3.0 and SSIS

Jul'17 to Apr'19 with S.S. InfoTech, MP as Software Developer

Skills - Python, MYSQL, PL/SQL, MS SQL Server, and ASP.NET

PROJECTS UNDERTAKEN

Organization	Affine Analytics
Client	SEBI
Domain	Securities Market Operations
Primary Goal	Data comes from the 6 types of exchanges and 13 types of segments wise in the form of CSV, applied all the data cleaning process before load data into the hive table, and after that added some business logic then Finally, data is loaded into the hive dimension table (historical records).
Solution	Post building it, with PySpark that took only 2 hours for all segments of NSE. Compared to using informatica earlier that took more than 12 hours

Organization	BSP Technologies
Client	Alder
Domain	Hospital Software
Primary Goal	Created end-to-end Big Data hospital application for 6 to 8 modules likes - patient registration, employee registration, patient appointment, role wise permission, patient Check-in, EHR (Electronic Health Records), and billing modules
Solution	Developed web application with technologies like - big data tools PySpark, Hive, Sqoop, SQL Server, PL/SQL, and AWS Clouds (S3, EC2), Airflow

Organization	S.S. InfoTech
Client	Yadi
Domain	Retail ERP
Primary Goal	Purchase, sale, buyer seller payment and stock maintenance record and reports modules
Solution	Developed web applications by using technologies like - Python, MYSQL, PL/SQL, MS SQL Server, and ASP.NET

CERTIFICATIONS

- ❖ Data science and data engineer master’s program from Simplilearn, 2022

EDUCATION

- ❖ B.E in Computer Science from RGPV University, Bhopal, 2016

ACHIEVEMENTS

- ❖ Ranked amongst ten thousand candidates in Pre-Engineering Test in MP

STRENGTH

- ❖ My biggest strength is my **positive attitude**. I know how stressful this industry can get, but I'm able to keep my calm.

PERSONAL DETAILS

Date of Birth: 2nd March 1994 | **Languages:** English and Hindi | **Address:** House No. - 1683 Gorakhpur Gurudwara Jabalpur (M.P) - 482001