



# Sagar Patro


Big Data Developer  
Mob: +91-8984129464



[sivasagar.patro@gmail.com](mailto:sivasagar.patro@gmail.com) 

[linkedin.com/in/sagar-patro-b8ba4613b](https://linkedin.com/in/sagar-patro-b8ba4613b) 

[stackoverflow.com/users/10172505/sagar-patro](https://stackoverflow.com/users/10172505/sagar-patro) 

Bangalore, India 

**Experience – 4.1 Years**

**Serving notice period - 26<sup>th</sup> October LWD**

**Working as Data Practitioner with Brillio**

---

## Professional Summary:

Having 4 years of experience in developing Big Data Applications using Azure cloud components and Hadoop Ecosystem Tools using Spark/PySpark and Hadoop technologies with end to end project implementation experience.

---

## Skills Set:

- |                      |              |
|----------------------|--------------|
| • Azure Databricks   | • Scala      |
| • Azure Data Factory | • Spark SQL  |
| • Azure DevOps       | • Hive       |
| • PySpark\Spark      | • Sqoop      |
| • Snowflake          | • Oracle     |
| • Python             | • SQL Server |
| • Hadoop             |              |

---

## Tools:

Sqoop, Databricks, SAP HANA, Informatica, Tidal, Crontab, JIRA, Service Now, Eclipse, SBT, Azure DevOps CI CD

---

## Work History:

### Accenture

May 2019- August 2020

Big Data Developer

**Client:** Devon Energy

**Project 1#** Replacement of SAP HANA by PySpark using Azure Databricks (End to End implementation)

- POC of project architecture for implementation/migration of SAP HANA's ETL logic to PySpark ELT code using Azure Databricks from scratch.
- Whole application development done using PySpark dataframe libraries and core Python in Azure Databricks notebooks to deal with SAP HANA ETL mappings implementation.
- Tuning spark code for better performance/optimization.
- Created Data Pipelines using Azure Data factory, keeping them task basis execution and monitoring them after production deployment.
- Deployment of Codes/Notebooks were done using Azure DevOps.

- Code releases from DEV to QA and from QA to PROD environment done using Azure DevOps following CI/CD process implementation.
- Codes were pushed and maintained in Azure Repos (Clubbed with GIT).
- Done data modelling for implementation of Snowflake database warehouse for the data migration from Oracle, SQL Server, ECC, MDM, TOW databases to Snowflake (Azure).
- Used Snowflake as Cloud Virtual Data warehouse, this acted as only source for all kinds of data for Databricks ELT implementation.

**Client:** Devon Energy

**Project 2#** Replacement of Informatica **by** PySpark using Azure Databricks (End to End implementation)

- Understanding the Informatica ETL mappings and designing the coding approach in PySpark and analysing the source data from Snowflake database.
- Implementation of same business logic in PySpark ELT methodology by referring the present critical mappings as designed in Informatica.
- Converted the Informatica ETL Mappings into PySpark dataframes and core Python using Azure Databricks.
- Storing the final Spark dataframe in a staging area in Snowflake database for downstream applications such as SAP HANA.
- Designed the Data pipelines using Azure Data Factory and keeping them execution as per task basis.

**HCL Technologies**

August 2016- May 2019

Hadoop Developer

**Client:** Clorox Company

**Project 3#** Retails Analytics & Planning

- Part of data modelling team to design the project flow.
- Data profiling for new project requirements
- Involved in Data Ingestion from various relational database providers on HDFS for analysis and other big data operations.
- Data transformation and data preparation for results using Spark API's, Dataframes, Datasets and HQL.
- Involved in creating Hive tables, loading and analysing data using hive queries.
- Created internal and external tables with properly defined static and dynamic partitions for efficiency.
- Implemented Partitions and Bucketing on Hive tables for efficient result query.
- Data migration from RDMS databases to HDFS.

**Client:** Iron Mountain

**Project 4#** Parts and Services

- Developed HSQL for creating data marts for data transformation.
- Involved in data transformation using Spark components to analyse and process the data.
- Extensively worked with Scala and Spark Data frames.
- Data Ingestion from various RDMS databases by using Sqoop.
- Developed Spark scripts by using Intillij as per the requirement

---

## Accomplishment:

- Got “Star award” by Accenture for code design and code delivery within the sprint SLA.
  - Got Pat on back (Internal Reward) for strategy implementation in Migration project.
  - Awarded two times with HCL Comnet silver jubilee award for outstanding performer.
  - Awarded two times best operation excellence in Clorox project.
  - Awarded best operation excellence from Iron Mountain client.
- 

## Education:

Bachelor of Technology Gandhi institute Of Engineering and technology, Gunupur, Odisha Electronics and Communication Engineering	July 2012- June 2016
Higher Secondary Education Lingaraj Science College, Brahampur, Odisha +2 Science, CHSE board	June 2010- May 2012
Matriculation DAV Public School, Brahampur, Odisha 10 <sup>th</sup> , CBSE Board	Jan 2000- May 2010

---

## Personal Details

- Name : Sagar Patro
- Date of Birth : 23.10.1993
- Languages Known : English, Hindi, Odia, Telugu

## Declaration:

I hereby declare that the above information is true to the best of my knowledge.