

Prashant Nagpal

Experience: 1.9 year

prashantnagpal.nagpal@gmail.com

+91 8950234707

Summary

High-Energy passion-driven individuals willing to use my skills and proficiencies to deliver value products for the end-user.

Skills

- Having 1.9 year of experience in the design and implementation of highly efficient and optimized ETL pipelines.
- Having good interpersonal skills, interaction with the team, ability to work independently.
- Experience in requirement gathering, infra-setup for building ETL pipelines(Open source).
- Experience in building data integration and data processing pipelines using open-source platform for batch and real-time, structured and unstructured data.

Additional Skills

- Qlik Replicate
- Qlik Compose for data warehouse

Work History

Data Engineer, Nov'19 to current

Company-Team Computers, Gurugram

Project Details:

1. Automobiles - Data Lake & DatawareHouse

Description: Designed and implemented the data lake in the Hadoop ecosystem defining 3 layers(Staging, Validated, Refined) and Data Warehouse in PostgreSQL. Therefore Preparing Dashboard in Qlik Sense.

Some of the multiple aspects of projects, which are:

- Implementation of data lake using HDFS.
- Developed data integration pipelines for Full-Load from source(Postgres, Flat files) to storage(HDFS)using Apache Spark and did Data Compression up to 70% (from CSV to Parquet).

- Developed ETL pipelines for data validation and refined layers using pyspark ETL and query the results using Apache Drill.
- Design and Developed a Data-Governance table in Postgres storing meta-data Information of all 3 layers.

Tools and Technologies:

Hadoop, Apache Spark, Python, Apache Drill, Shell-Scripting

2. Insurance Data lake

Description: Developed and Implemented ETL jobs for data ingestion and applying transformation on source data by using big data frameworks like apache spark, apache hadoop.

Involved in multiple aspects of projects, which are:

- Understanding the business requirement and preparing a data dictionary.
- Implementation of legacy ETL in spark with optimization.
- Worked on performance tuning of spark jobs.
- Installation and infra setup of Hadoop and spark cluster.

Tools and Technologies:

Apache Spark, Apache Hadoop, Apache Thrift, PostgreSQL, dbeaver.

Qualifications

Degree	Institute	Year	Percentage
B.Tech	Panipat Institute of Engineering & Technology, Panipat	2019	71%
12th	Pratap Public School, Karnal	2015	61 %
10th	Pratap Public School, Karnal	2013	7.2 C.G.P.A