# Ashirvad Kumar

Bangalore, India 🏠
(+91) 8877284267 📞
Ashirvadkumar.py@gmail.com ✉
/in/ashirvadkumar/ 🔗

**Data Engineer** with 3+ years of experience in architecting, creating, testing and maintaining the complete data warehousing and business intelligence solutions, management and processing systems. Worked with top global insurance clients with exposure to technologies and tools like AWS , Redshift, Glue. Informatica, Hive, Python scripting, MPP, Git and Team city.

## Skills

• AWS •Redshift Glue•AWS S3• Informatica • Spark• IBM Data Studio • SQL, NoSQL • Shell scripting • Python• Unix• PL/SQL• Stored Procedures• Data Warehousing• Star/Snowflake schemas• Data Modelling• Schema design• CA Workload Automation.

## Ernst & Young,  Data Engineer
## Experience  (Feb 21 Till date)

FEBRUARY 2021 – PRESENT
**Role**: - Data Engineer/ Retail

**Project Description: the project involves gathering data from old legacy systems and modernizing them to new frameworks including spark, python and AWS platform.**

**Key Metrics:-**

- Analyzed existing SSIS packages and extracted ETL information to moderate the same on snowflake
- Created DAGs to automate the data flow
- Used SQL based ETL to maintain the data flow
- Created views to refresh the dashboards.

## Cognizant,  Data Engineer
## Experience  (2.11 years)

# Projects

DECEMBER 2019 – PRESENT
**Role**: - Data Engineer/ Global Insurance Client
**Project**: - New York Life.

**Project Description** - New York life involves producing Policy and Coverage information of consumers on monthly basis. This process includes extraction of data from data lake (RIW), comparing the matched policies with Data Warehouse (ODS) and calculating the amount and other benefits for new users..

**Technologies used**: - AWS, Redshift, S3, Hadoop, Spark, EMR

**Key Responsibilities:**

- Design and Develop ETL Processes in AWS Glue to migrate Campaign data from external sources like S3, ORC/Parquet/Text Files into AWS Redshift.
- Maintained the data pipeline from source (e.g., AWS SNS, SQS, Kinesis, S3) to S3 bucket, finally to Redshift using SQL.
- Guaranteed the data integrity by creating python package to triggered alarm if finding missing files in S3 bucket.
- Data Extraction, aggregations and consolidation of Adobe data within AWS Glue using PySpark.
- Performed Data Migration from Amazon S3 to Amazon Redshift.
- Designed External and Managed tables in Hive and processed data to the HDFS using Sqoop.
- Created user defined functions UDF in Redshift.
- Provided QA support, deployment and reporting using Unix scripts.
  .

FEB 2019 – NOV 2019
**Role**: - Data Engineer/ Global Insurance client
**Project**: - @LAST Data Remediation.

**Project Description** -. @LAST (Life Amenities System Transformation) was created to consolidate the existing 15 ODS sources into one source named VpasOne.

**Technologies used**: Hadoop, MapReduce, HDFS, Hive

**Key Responsibilities:**

- Worked on analyzing Hadoop cluster and other big data analytics tools such as Pig, hive
- Responsible for building scalable distributed data solutions using hadoop.
- Involved in exporting/importing data from local/external file system and RDBMS to HDFS
- Used PIG for ETL data cleansing, integration and transformation.
- Exported and analyzed data to relational databases using sqoop for visualization and reporting's.
- Created smooth procedures to reflect batch reporting based on source systems.
- Created and maintained technical documentation for launching Hadoop clusters and maintain hive queries.

MARCH 2018 – FEB 2019
**Role**: - ETL Engineer / Global Insurance client
**Project**: - JHIM Operational data source.

**Project Description** -. JHIM ODS was designed to integrate data from multiple sources for additional operations on the data, for reporting, controls and operational decision support. It includes data from 15 regions across United states. The processes involved creating daily/weekly/monthly jobs to load data on tables based upon frequency.

**Technologies used**: - Hive, Spark, Informatica, SQL, CAWLA, Putty

**Key Responsibilities:**

- Created process to load data to NLZ using UDE, XML, CSV, Flat files which are FTPed on server.
- Created truncate load NLZ jobs scheduled to run daily.
- Used data from NLZ layer to load into ODS based on source systems and requirements
- Worked with Product owners, stake holders, business analysts to understand business process.
- Designed schema to load policy and coverage level information following the methods of snowflake.
- Create temporal tables to keep history data based on weekdays.
- Created SCD type 2 jobs to enhance data usage for OLAP
- Created monthly, weekly and daily load jobs.
- Created Stored Procedures to load and insert policy and coverage information to multiple tables.

# Education

AUG 2017

### B.Tech in Computer Science/ Netaji Subhas Institute of Technology, Patna
1. Created Android Application for college fest in the year 2016 and 17.
2. Worked on Home Automation system as a part of final year project.

### HSC/ Patna Central School, Patna
1. 12th standard with Physics, chemistry and Mathematics as a core subject from CBSE.
2. Appointed house caption for two years.

### SSC/ Patna Central School, Patna
1. 10th Standard with Science and Mathematics as a core subject from CBSE.

# Achievements

- Awarded as **Technocrat** in the 3rd quarter of 2019 for excellent deliverables.
- Awarded as **Shining star** in the 1st quarter of 2019 for excellent deliverables.
- **Microsoft Student Partner :** As a part of MSP India 2016 program, I was selected as **MSP** and was working as a campus leader for Microsoft.
- **Campus Ambassador** by **geeksforgeeks.org** (2016-17)