Chhayansh Purohit

Dallas, TX | (945) 217-7878 | chhayanshpurohit@gmail.com| https://www.linkedin.com/in/Chhayanshp11/

EDUCATION

The University of Texas at Dallas **Aug 2022-May 2024**
Master of Science, Information Technology and Management GPA: 3.86/4.0
Graduate Teaching Assistant
Relevant Coursework: Big Data, Data Visualization, Business Data Warehousing, Advanced Statistics for Data Science

Jabalpur Engineering College, India **Aug 2017-May 2021**
Bachelor of Technology, Computer Science and Engineering GPA: 3.7/4.0
Relevant Coursework: Data Structure and Algorithms, Database Management System, Software Engineering, Data Mining

SKILLS

Certifications	AWS Cloud Practitioner, Alteryx Certified Core Designer
Programming	Python (NumPy, Pandas), SQL, C, C++, R, Unix
Database	MySQL, PostgreSQL, SQL Server, AWS (Redshift, DynamoDB, S3)
Cloud	AWS (S3, EC2, RDS, Redshift, VPC, Glue, Athena, Lambda, Quick Sight), GCP, Big Query
Data engineering	ETL Pipeline, ELT, Kafka, Airflow, Spark, Hadoop (HDFS, Sqoop, Hive), Snowflake, Docker, Tableau

WORK EXPERIENCE

Data Engineer 2 | LPL Financial | Austin, TX **June 2024–Present**

- Successfully designed and setup **AWS Glue** ETL pipeline to automate data processing tasks using **Lambda** functions, which significantly reduced manual intervention, enhancing data quality and consistency.
- Orchestrated complex data transformation process utilizing **AWS Step Functions** and **Airflow Dag's** ensuring smooth execution of Glue ETL jobs for further data refinement.
- Applied advanced **dimensional modeling** techniques creating fact and dimension tables, to represent complex financial data using PostgreSQL in a **Snowflake** environment, enabling easy navigation and reporting.

Data Engineering Intern | LPL Financial | Austin, TX **May 2023–May 2024**

- Developed update notification **API** to ensure precise dissemination of notifications for **ETL** job monitoring and alerts.
- Developed and executed a comprehensive data cleansing project using **SQL**, **NumPy** and **Pandas**, resulting in a 30% reduction in data errors and improving the accuracy of forecasting models.
- Developed a **Docker** Image optimized for unit testing purpose resulting in increased efficiency and code quality by 25%

Data Engineer | Infosys Limited | India **Sep 2021–Jun 2022**

- Developed and implemented a Rules Engine for a leading healthcare company, resulting in 60% increase in efficiency for their Registry Quality Control Management System by capturing changed data and conducting logic rule checks using SQL.
- Designed and executed workflows and **SQL** queries to clean and transform raw clinical data from global sources, ensuring accurate and reliable information for analysis.
- Enhanced data management and reporting capability by automating routine **ETL** tasks by designing **Alteryx Workflows** and **Batch Macros**, reducing manual operations by up to 35%.
- Proactively monitored **Data Pipelines** and promptly resolved data quality issues, resulting in a 25% decrease in data-related incidents and ensuring data integrity and consistency throughout the pipeline.

ACADEMIC PROJECT EXPERIENCE

End to End Data Engineering on AWS May 2023 – Jun 2023

- Spearheaded a YouTube video analysis project, leveraging Python and AWS services (S3, Glue, Lambda) to extract, transform and load the data in Redshift warehouse empowering real-time analytics with 50% quicker responses using Athena.
- Automated the pipeline with Airflow DAG and used Quick Sight to visualize video trends based on category and location.

Truck Fleet Risk Analysis using Sensor Data Mar 2023 – May 2023

- Ingested 10000 records to HDFS with Sqoop and analyzed using Hive and Pig to find top 5 risky drivers and over-used trucks.
- Integrated HDFS with Tableau for accidental trends, forecasting vehicle maintenance and driver safety with 92% accuracy.

PUBLICATIONS

Trusted Infrastructure Design for Secure Virtualization in Cloud Computing Dec 2021–Sep 2022
An Informative Analysis of Encryption Algorithms Using Quantitative Fitness Function | Springer Jul 2019–Dec 2019