

Chhayansh Purohit

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

EDUCATION

The University of Texas at Dallas

Master of Science, Information Technology and Management

Relevant Coursework: Big Data, Data Visualization, Business Data Warehousing, Advanced Statistics for Data Science

Aug 2022-May 2024

GPA: 3.86/4.0

Jabalpur Engineering College, India

Bachelor of Technology, Computer Science and Engineering

Relevant Coursework: Data Structure and Algorithms, Database Management System, Software Engineering, Data Mining

Aug 2017-May 2021

GPA: 3.7/4.0

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

- Engineered AWS Glue-based ETL pipelines with automated Lambda triggers, minimizing manual effort and elevating data integrity across workflows.
- Optimized data ingestion by enabling direct vendor-to-cloud transfers, cutting latency and streamlining real-time data accessibility.
- Developed robust data validation frameworks, implementing proactive error handling measures that reduced discrepancies by 15% and ensured seamless ETL operations.
- Spearheaded the integration of a new data provider, architecting an automated pipeline with AWS Lambda and validation mechanisms to facilitate error-free data ingestion.
- Managed the release and deployment of the enterprise data integration framework, overseeing pre and post-release coordination to ensure smooth transitions and minimal disruptions.

Data Engineering Intern | LPL Financial | Austin, TX

Aug 2023–May 2024

- Collaborated with cross-functional teams to deliver high-quality data products in an agile development environment.
- Worked with both on-premises and cloud-based database and data warehouse solutions, including Oracle, SQL Server, and AWS.
- 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** package optimized for unit testing purpose resulting in increased efficiency and code quality by 25%

IT Audit Information Technology Intern | LPL Financial | Austin, TX

May 2023–Aug 2023

- Performed walkthroughs and testing of IT Controls (CRTP), SOC1 reports and FICCA compliance.
- Tested and documented work performed, communicate requirements and follow up with the stake holders and control owners across Technology, Finance and Business organizations in a concise and consistent manner.
- Participated and supported in ad hoc audits such as AWS Cloud systems as part of a Cloud Audit, Third-party vendor audit, Data Security, ensuring adherence to compliance requirements, internal standards, and industry best practices.

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%.

Software Engineer 1 | Worldsoft IT Solutions Pvt Ltd | India

Jan 2020–July 2021

- Created a Python-based ETL (Extract, Transform, Load) workflow to automate data retrieval from public APIs while streamlining data cleaning and transformations, resulting in a 30% reduction in data processing time.
- Devised an in-house framework to export on-demand Google Analytics data into the SQL Server, reducing manual effort by 40% and eliminating the need for external tools. Implemented intuitive UIs using Angular, HTML, CSS, and Python.
- Developed a DFS-based PDF parsing algorithm, enabling the extraction of financial data from complex PDF files. This automation reduced data extraction time by 50%, enhancing data processing efficiency.

Cloud Engineer Associate | Worldsoft Technologies Pvt Ltd | India**Jan 2019–Sept 2019**

- Devised a scalable and distributed data infrastructure in Spark and Python to optimize ELT pipelines, enforce quality checks, apply business rules, and store in a data warehouse, ensuring data availability to downstream apps with 95 - 100% accuracy.
- Formulated unit tests for PySpark code using Pytest, ensuring 95% code reliability by detecting bugs early on.
- Collaborated with agile teams to gather requirements, debug, and resolve critical issues, participate in code reviews, develop reusable code, and document best practices, improving team productivity and reducing development time by 20%.

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****LEADERSHIP AND ACHIEVEMENTS**

- Graduate Teaching Assistant – UT Dallas Jan 2023 – May 2023
- Student Assistant – UT Dallas Jan 2022 – Jan 2023
- Product Base Club – UT Dallas Aug 2022 - Present
- Infinity Loins Club – UT Dallas Aug 2022 - Present
- General Secretary - ACM Student Chapter, JEC Jabalpur Nov 2019 – Nov 2020