

ARPITA DESHMUKH

Dallas, TX | H1-B Visa (Active from Oct 2021)

deshmukharpita728@gmail.com | arpitapd@gmail.com | [linkedin.com/in/arpitapd](https://www.linkedin.com/in/arpitapd)

ABOUT ME

I'm a results-driven data engineer with 4+ years' experience in building & automating ETL pipelines (Python, Alteryx), data modeling, analyzing & integrating large data sets with millions of rows (SQL), and creating dynamic, high-data quality dashboards (PowerBI). Proven ability to streamline data processes and improve operational efficiency

SKILLS

Languages & Skills: SQL | Python | JSON | Agile | ETL pipelines | Git | Data Modeling | Data warehouse | API

Tools: Alteryx | PowerBI (DAX, M query, data visualization & data modeling) | RDBMS | SQL Server | SSIS | SSRS | SSAS | AWS (S3, EC2, Redshift, AWS Athena) | Airflow | Spark

Other skills: Neo4j | NoSQL (MongoDB) | ML & Data Mining (Beginner level)

EXPERIENCE

Behavioral Health Group, LLC (BI Data Engineer)

Dallas, TX Jan 2022 – Present

- Collaborated with C-suite executives to capture requirements, identify relevant **KPIs** & trends, leading to data-driven decisions across for 120+ global locations by delivering end-to-end data solutions from ETL to reporting
- Built and maintained automated **ETL** data pipelines using **Alteryx, SQL**, integrate data from multiple data sources; built insightful **PowerBI** reports, improving operational decisions significantly
- Streamlined **JSON** data extraction from call center **APIs** with **Python ETL**, making process **100%** automated
- Automated **API** data extraction, transformations in **Python**, backing up and uploading report to HR system via SMTP using **Alteryx** and **batch** scripting, eliminating manual efforts completely
- Created & optimized **SQL** queries using complex joins and window functions, improving data retrieval by **15%**
- Implemented Row Level Security (**RLS**) in **PowerBI**, enabling tailored data access based on employee location and job title, providing personalized, role specific insights for different user groups

Behavioral Health Group, LLC (Revenue Data Analyst)

Dallas, TX July 2020 – Jan 2022

- Expedited **ETL pipeline** by replacing normal functionality to In-DB process in Alteryx, reducing execution time by **80%**
- **Patient No Show Project** - Leveraged **SQL** (CTEs, window functions, joins) to analyze over **7M+ records** and created PowerBI reports to track missed doses and follow-up calls, enhancing patient care management.

Strategic Materials, Inc (Data & Business Intelligence Analyst Intern)

Houston, TX May 2019 – Aug 2019

- Conducted requirements elicitation from C-suite executives, and stakeholders for creating interactive dashboards.
- **Tracking on-Time shipments** - Created BI dashboard to compare on-time shipments percentage across global facilities; incorporated email alert system if on-time percentage falls below 70% for the current month.

PROJECTS

Youtube Data Analysis

[AWS S3](#) | [Glue](#) | [Athena](#) | [Lambda](#) | [IAM](#) | [Quicksight](#) | [ETL](#) | [SQL](#)

- Built an end-to-end pipeline using AWS services to analyze YouTube data, automating ingestion of JSON and CSV data files, transformations, querying in Athena and reporting on Quicksight

Movie Recommendation System

[Graph Database - Neo4j](#) | [MariaDB with Galera clustering on AWS](#) | [SQL](#) | [Cypher](#)

- Processed 3GB of Netflix and IMDB data to enable users to search for movies, view friends' watchlists, and receive personalized recommendations based on ratings and viewing history

EDUCATION

• **Texas A&M University** (MS in Management Information Systems)

College Station, TX Aug 2018 - May 2020

• **University of Pune** (Bachelor's in Information Technology)

Pune, India Aug 2014 - May 2018