ARPITA DESHMUKH
Data Engineer | Storytelling with Data

✉ arpitapdtamu@gmail.com 🔗 https://arpitapd.github.io/ in linkedin.com/in/arpitapd (Location: Dallas, TX)

Summary

Data engineer with 5+ years of experience transforming complex datasets into business insights and delivering end to end analytics projects. Skilled in designing automated ELT/ETL data pipelines (Python, Alteryx), analyzing large datasets (SQL), and creating executive-ready dashboards (Power BI). Partnered with stakeholders to identify opportunities, understand business needs and deliver data insights to enable data-driven decisions.

Skills

Programming & Querying: SQL, Python, R, DAX
Data Visualization: Power BI, Looker, Tableau (familiar)
Tools: Alteryx, RDBMS (SQL Server), AWS (S3, Glue, Redshift, Athena), Airflow, BigQuery, GCP
Other skills: ETL, Data Modeling, Data Visualization, Quantitative Analysis, API Data Extraction, Recruiting Metrics

Work Experience

BI Data Engineer   **Jan 2022 – Present**
Behavioral Health Group   *Tools Used: SQL, APIs, Python, Alteryx, Power BI*

- Developed **dashboards** that streamlined **KPI tracking** across 120+ global locations to aid strategic decision-making
- Evaluated **15 million+ rows** using SQL, PowerBI to identify **missed call patterns**, enabling **improved staffing decisions**, reducing missed calls from **22% to 14%**
- Created optimized SQL queries using complex joins and window functions, **improving data retrieval speed by 15%**

People Analytics Projects
Behavioral Health Group

- Automated scalable Python pipeline to extract **employee Roster and Termination data** from HR system into Azure, **saving 10 hours** of manual reporting per week
- Developed dashboards to track **recruiting metrics**- headcount, turnover, new hires, average days positions are open, average tenure per position, providing leadership with actionable insights for workforce planning
- Implemented Row Level Security (RLS) in Power BI to **protect sensitive people data**, enabling data access based on job title
- Deployed Python pipeline including API data extraction of **people data**, transformations in Python, backing up and uploading report to HR system via SMTP using Alteryx and batch scripting, **saving $10K in labor costs**

Revenue Data Analyst   **Jul 2020 – Jan 2022**
Behavioral Health Group   *Tools Used: SQL, Python, Alteryx, Power BI*

- Analyzed **40 million+** records using SQL (CTEs, window functions, joins) to track missed doses and follow-up calls; reducing patient no-shows by 10% in 3 months, **estimated $300K in recovered revenue**

Business Intelligence Analyst Intern   **May 2019 – Aug 2019**
Strategic Materials, Inc   *Tools Used: Data Modeling and Visualization in Targit*

- Partnered with C-suite to define KPIs, built dashboards to track daily-to-yearly sales, operations-focused metrics for 50 plant locations; delivering BI projects **4 months before deadlines**

Projects

YouTube Data Analysis in AWS   **AWS S3, Glue, Athena, Lambda, IAM, Quicksight, ETL, SQL**

- Implemented serverless ETL pipeline processing 200+ daily trending videos across 10 regions; analysis revealed 30% videos reached trending within 24 hours, highlighting **early engagement is key for virality**
- Identified 60% of videos trending in Japan were Animation, compared to 15% in US, indicating **region-specific content strategies**

Education

Masters in Management Information Systems   **Aug 2018 – May 2020**
Texas A&M University   *College Station, Texas*

Bachelors in Information Technology   **Aug 2014 – May 2018**
Pune University   *Pune, India*