ARPITA DESHMUKH
Data Engineer | Business Analyst

✉ 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 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). Hands-on experience with cloud platforms (AWS, GCP) through personal projects. Partnered with stakeholders to identify opportunities and business needs, and communicated data insights clearly to enable data-driven decisions

Skills

Programming & Querying: SQL, Python, R, DAX
Data Visualization: PowerBI, Looker, Tableau (familiar)
Tools: Alteryx, SQL Server, AWS (S3, Glue, Redshift, Athena), Airflow, BigQuery, GCP
Other skills: ETL, Data Modeling, Data Visualization, Quantitative Analysis, API Data Extraction, Statistical Analysis, Machine Learning Concepts

Work Experience

BI Data Engineer   **Jan 2022 − Present**
Behavioral Health Group   *Tools Used: SQL, APIs, Python, Alteryx, Power BI*
   − Developed **executive-ready dashboards** that streamlined KPI tracking across 120+ global locations
   − Automated Call Center JSON data extraction via APIs using Python, **saving manual efforts by 10 hrs/week**
   − Evaluated **15 million+ rows** using SQL, PowerBI to identify missed call patterns, enabling improved staffing decisions, **reducing missed calls from 22% to 14%**
   − Deployed Python pipeline including API data extraction, transformations in Python, backing up and uploading report to HR system via SMTP using Alteryx and batch scripting, making process 100% automated, **saving $10K in labor costs**
   − Created optimized SQL queries using complex joins and window functions, **improving data retrieval speed by 15%**

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**
   − Optimized Alteryx ETL flows using In-DB processing, **reducing runtime by 80%** for timely clinical reporting and improving operational efficiency

Business Intelligence Analyst Intern   **Jul 2020 − Jan 2022**
Strategic Materials, Inc   *Tools Used: Data Modeling and Visualization in Targit*
   − Partnered with C-suites to define KPIs, create dynamic dashboards tracking daily-to-yearly sales, operations-focused metrics for 50 plant locations globally; delivering BI projects **4 months before deadlines**
   − **Tracking On-Time Shipments** – Built BI dashboard to track on-time shipments across global facilities, implementing automated alerts for <70% on-time rates, **enabling quick response and avoiding delays**

Projects

YouTube Data Analysis in AWS   **AWS S3, Glue, Athena, Lambda, IAM, Quicksight, ETL, SQL**
   − Implemented a serverless ETL pipeline processing 200+ daily trending videos across 10 regions; analysis revealed 30% reached trending within 24 hours, highlighting **early engagement is key for virality**
   − Identified that 60% of videos trending in Japan were Animation, compared to 15% in US, indicating **region-specific content strategies**

ELT Pipeline using GCP (Mini Project)   **GCP, BigQuery, Airflow DAGs, Looker**
   − Built an ELT pipeline on GCP using BigQuery and Airflow DAGs to process 1 million+ records, created a Looker dashboard with summary analytics; gaining hands-on experience in end-to-end cloud analytics workflows

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*