Mayuresh Deodhar

Brookline MA 02445 | 857-869-2403 | deodharmayuresh99@gmail.com

www.linkedin.com/in/mayureshd | https://github.com/Mayuresh-Deodhar

EDUCATION

Master of Science in Information Systems, Northeastern University, Boston, MA   **May 2020**

- **Relevant Courses***: Designing Data Architecture for Business Intelligence, Data Science Engineering Methods and Tools, Data Management and Database Design

Bachelor of Engineering in Information Technology, University of Mumbai, Maharashtra, India   **May 2018**

- **Relevant Courses***: Object Oriented Programming Methodology, Data Mining and Business Intelligence

TECHNICAL SKILLS

Programming Languages:	SQL, Python, Pyspark, Spark SQL, Java
Databases:	Snowflake, Redshift, MySQL, Oracle 11g, Microsoft SQL Server, PostgreSQL, BigQuery
Business Intelligence Tools:	Tableau, Microsoft Power BI, Looker
ETL Tools:	Talend Studio, Alteryx, SSIS
AWS tools:	EC2, Glue, S3, Lambda, Step Functions
Tools:	Databricks, GitHub, Airflow

WORK EXPERIENCE

Verana Health   **Boston, USA**

Data Engineer   **August 2021 – April 2024**

- Analyzed various EHR systems in AWS Athena identifying source to target mappings to build a sustainable central repository
- Applied and maintained complex data transformations to mapping scripts based on business logic and requirements
- Parallelized Databricks jobs using airflow to build scalable ETL pipelines (pyspark) ingesting data for over 1500 practices into datalake capable of supporting broad range of use cases in analytics, real world evidence, clinical trials and MIPS reporting
- Performed data migration testing, quantitative and qualitative analysis to verify data integrity and reliability
- Collaborated with clinical and analytics teams to enhance the quality of ingested data by 35%, leveraging a deep understanding of clinical requirements and KPI metrics.

Massachusetts General Hospital   **Boston, USA**

Data Engineer   **November 2020 – August 2021**

- Utilized and analyzed EHR/EMR using Snowflake to understand the risk factors and mechanisms of Alzheimer's disease
- Constructed ETL pipelines in python to extract data from EDW, transform it into required metrics and load them in flat files
- Collaborated with researchers/physicians to identify goals of the research and presented summary statistics using Tableau
- Processed and validated large datasets using regular expressions in python to reduce false positives by 25%

Upped Events   **Boston, USA**

Data Analyst   **July 2020 – October 2020**

- Analyzed KPIs and trends in Tableau and SQL to improve events revenue by evaluating opportunities in the existing events
- Developed ETL jobs on Events and product data using SSIS to turn disorganized data feeds into high-quality data assets
- Communicated with stakeholders to understand evolving business requirements and effectively managed end to end impact analysis and component changes
- Worked with CEO/advisors to ideate, design, and integrate company's client-facing analytics report -- synced in real-time to the company's database, programmatically changing based on an event's chosen features

Entel Networking Systems   **Mumbai, India**

Data Analyst (Internship)   **August 2017 - August 2018**

- Translated operational metric requirements into functional requirements by constructing complex SQL queries
- Built SSIS ETL Packages to integrate sales and client data from flat files into SQL Server to construct data marts
- Enabled Slowly changing dimensions to track the client and sales information enabling historical analysis
- Automated KPI reports by deriving insights through dynamic visualizations in Power BI

ACADEMIC PROJECTS

Retail Data warehouse (Tableau, Power BI, ETL, Talend, SSIS, ER Studio, Data modeling, Data mart)

- Designed and developed product-based company data warehouse star schema dimensional model in ER studio
- Performed transformations and integrated data from heterogeneous sources – SQL server, MySQL, Oracle, PostgreSQL, CSVs and Implemented error handling, currency conversion and slowly changing dimensions in Talend and SSIS
- Incorporated parallelization to maximize job performance by reducing execution time by 25%
- Generated custom analytical dashboards to pull insights on sales and inventory in Tableau and Power BI

Recommendation System (Python, Flask, AWS, S3, EC2, Tableau)

- Developed a flask application for recommender system and RFM modeling-based customer segmentation
- Deployed the application on AWS EC2 instance and created analytical dashboards in tableau public for marketing insights
- Optimized the model hyperparameters to achieve a rank of 57 from 920 and AUC metric of 0.862 using Grid Search CV
- Created joins, subqueries, stored procedures, and triggers to retrieve components at various locations on ISS