Ieshaan Sharmaa

iesharma196@gmail.com • (716) 507-9080 • https://www.linkedin.com/in/ieshaan/• https://github.com/Ieshu

Summary

Experienced Data Engineer proficient in data pipelines, analytical modeling, and data warehousing optimization. Skilled in ETL/ELT processes, data modeling, and BI tools like Looker/Tableau. Expertise in leveraging dbt and Snowflake to drive actionable insights.

Professional Experience

Data Engineer II, ITVorks, Seattle, WA   Aug 2023 - Present
- Implemented AWS Glue for ETL pipelines, resulting in a **30% decrease in data ingestion latency**, enhancing real-time analytics capabilities for critical patient data.
- Capitalized **AWS S3 lifecycle policies coupled with Glacier storage** for archiving historical patient records. **Reduced storage costs by 40%** while ensuring compliance with data retention regulations.
- **Migrated legacy on-premises data processing workflows** to serverless architectures leveraging AWS Glue and AWS Lambda. **Achieved a 30% reduction in infrastructure costs** while maintaining scalability and reliability, enabling efficient utilization of healthcare company resources.
- Optimized **AWS Lambda functions** and **Amazon Kinesis Data Firehose** to **reduce the latency** of real-time patient monitoring **data ingestion by 40%.**

Data Analyst, Amazon Web Service, Seattle, WA   July 2022 – July 2023
- Conducted scalable data storage solutions in **Amazon S3**, accommodating a **50% increase in data volume** without a proportional rise in storage costs.
- Orchestrated a **robust patching solution** for AWS EC2 instances, resulting in **a 40% reduction in vulnerability exposure** by ensuring timely application of security patches across the entire EC2 fleet.
- Monitored and **audited SQL scripts & and jobs in DBT** to analyze the performance and efficiency of run time of the jobs, and recommended changes resulting in **performance improvement by 23%** (cost savings $118K annually).
- Utilized **AWS Athena for ad-hoc querying**, providing data analysts with quick insights and **reducing query response time by 35%.**

Data Engineer I, Peacock Solar, Gurgaon, India   Jan 2020 - Sept 2020
- Enhanced solar panel efficiency by 15% through data analysis and optimization of panel positioning and orientation. This resulted in an annual **energy generation increase of 5,000 MWh**.
- Designed an **ETL workflow with AWS Glue** to integrate data from multiple sources, **resulting in a 30% reduction** in data **integration time** and improving data availability for analysis by 25%.
- **Reduced data replication latency by 50%** through high-speed and secure cross-region data transfers **using AWS DataSync**. Enhanced data availability and disaster recovery capabilities, achieving a 99.99% data replication accuracy rate.
- **Deployed AWS IoT Analytics for detecting and diagnosing faults** in solar panel systems based on historical performance data and machine reduce **mean time to repair (MTTR) by 30%** through automated fault detection alerts and root cause analysis.

Data Engineer, BLP Industry.AI, Bengaluru, India   Jun 2018 - Jan 2020
- Implemented an AD Inverter Deterioration model for Tata Solar, **improving solar power production by 10% and achieving an AUC of ~0.83 on AWS EC2** using python libraries and PySpark.
- **Increased wind turbine efficiency by 12%** using DBSCAN clustering, resulting in an annual energy generation boost of 10,000 MWh.
- Optimized renewable energy forecasting models with Python, **reducing forecast error by 25%** and enabling better resource allocation, **saving ₹100,000 annually**.
- Conducted analysis on weather data and solar irradiance patterns, optimizing solar panel positioning and causing a **12% increase in energy generation and ₹500,000 annual revenue boost**.

Core Competencies
- **Programming**: Python, R, SQL, and C.
- **Visualization**: Looker, Tableau, Power BI, Qlik, Shiny, Adobe Analytics, Excel VBA, Chart.js Google Charts, Grafana
- **Technical:** Machine Learning, Data Analysis, A/B Testing, Apache Spark, Python, R, Data Science, Microsoft Excel, SQL, ETL Pipeline, MATLAB, Regression, Docker, Kubernetes, CI/CD, Kafka, Data Modelling(V-look Ups, Advanced Excel)
- **Database Systems:** Relational Database, PostgreSQL, MySQL, SQL, NoSQL, Database, MongoDB, Cassandra, Redis, Data Warehousing, Amazon Redshift, Google Big Query, Snowflake, IBM Db2
- **Big Data Technologies:** Hadoop, Map Reduce, Pig, Hive, DBT, DevOps, Terraform, Apache Airflow, Elasticsearch
- **Cloud Technologies:** S3, EC2, Lambda, AWS Glue, EMR, Dynamo DB, AWS Athena, Azure, Blob, GCP, Big Query, Big Table, Google Cloud Storage, HBase

Education

University at Buffalo, The State University of New York   Buffalo, NY
Master of Science in Engineering Science(Data Science)   Jan 2021

SRM University   Chennai, India
Bachelor of Technology in Computer Science Engineering   May 2018