

NIHARIKA PENDEM

Open to Relocation | niharikapendem57@gmail.com | (361) 688-1729 | [LinkedIn](#) | [GitHub](#)

SUMMARY

Microsoft Certified Fabric Data Engineer with 5+ years of experience designing and deploying data pipelines, analytics platforms, and cloud solutions across AWS, Azure, and GCP. Proficient in Python, SQL, Spark, Airflow, and dbt, with expertise in real-time streaming, data lake architecture, and cross-cloud ETL automation. Adept at building reliable data systems that drive decision-making and operational efficiency.

TECHNICAL SKILLS

Languages:	Python, SQL, Java, Shell Scripting, JavaScript, HTML, CSS
Cloud:	AWS (S3, Lambda, Redshift), Azure (Data Factory, Synapse), GCP (Dataflow)
Tools:	Git, Postman, VS Code, Eclipse, PyCharm, SQL Developer, Docker, Kubernetes, Jira, Github Actions
Libraries:	Pandas, NumPy, Scikit-learn, PyArrow, Spark, Kafka, Matplotlib, Seaborn
Databases & ETL:	Snowflake, BigQuery, PostgreSQL, MySQL, MongoDB, Apache Airflow, Glue, Databricks, dbt
Visualization:	Power BI, Tableau, Looker, Plotly, BIRT Reporting Tool

EXPERIENCE

Programmer Analyst	<i>Texas A&M University, Corpus Christi, TX</i>	Jan 2024 - May 2025
<ul style="list-style-type: none">Engineered automated data pipelines using Python, SQL, and Apache Airflow, ensuring accurate ingestion, transformation, and validation of academic and administrative datasets, reducing manual processing time by 60%.Integrated cloud-based ETL workflows with AWS S3, Snowflake, and Power BI to centralize interdepartmental data, streamline reporting workflows, and reduce data retrieval time by 6+ hours per week for university teams.Developed and deployed custom dashboards using Power BI, Tableau, and AWS QuickSight, delivering actionable insights to faculty and administrators for data-driven decision-making.Orchestrated batch job automation and scheduling using Airflow DAGs, enhancing operational efficiency by 50%, minimizing manual intervention, and enabling consistent data delivery across multiple university departments.		
Data Engineer	<i>Accenture, India</i>	Jan 2020 - Jul 2023
<ul style="list-style-type: none">Designed and deployed scalable ETL pipelines using Apache Spark and Azure Data Factory to automate ingestion of 10M+ daily smart meter readings into Snowflake and AWS Redshift, reducing data processing time by 30%.Automated billing and health monitoring reports leveraging Python, PySpark, and custom EnergyIP VEE rules, achieving 99.9% data accuracy and enhancing system reliability through early anomaly detection and validation controls.Optimized complex SQL queries and enhanced database performance in MySQL by applying indexing, partitioning, and query optimization techniques, reducing execution time by 40%.Created stakeholder-facing reports using BIRT and SQL-based tools to deliver actionable insights on meter exceptions, billing anomalies, and operational summaries under strict SLA timelines with executive visibility.Collaborated with cross-functional teams to support EnergyIP MDMS customization and upgrade initiatives, integrating Java-based logic, Kafka streams, and ETL validation checks for consistent and reliable meter data management.		

CERTIFICATIONS

• Microsoft Certified: Fabric Data Engineer Associate DP-700	Jun 2025 - Jun 2026
---	---------------------

EDUCATION

Texas A&M University, Corpus Christi, TX <i>Master of Science in Computer Science</i>	Aug 2023 - May 2025
--	---------------------

PROJECTS

YouTube Trending Data Analytics Pipeline using AWS	Sep 2024 - Oct 2024
<ul style="list-style-type: none">Launched an end-to-end data pipeline on AWS for analyzing YouTube trending videos, automating data ingestion, ETL and visualization. Built to identify content engagement patterns across regions, supporting digital marketing strategy.Used S3, Glue, and Lambda to process raw JSON data, transforming it into structured tables in Athena for querying.Crafted engaging QuickSight dashboards to track video engagement, category trends, and regional popularity.	
Amazon Web Scraping and ETL Pipeline	Jul 2024 - Aug 2024
<ul style="list-style-type: none">Implemented a Python-based web scraping pipeline using BeautifulSoup & Requests to extract product details (title, price, ratings) from Amazon. Designed to support competitive pricing analysis and track product availability trends.Structured and stored scraped data in CSV & PostgreSQL, optimizing scraping logic to handle pagination, dynamic content rendering and rate limits efficiently. Enabled seamless integration into analytics workflows.	

PYTHON PROJECTS

• Executed Python-based games including Snake, Connect 4, Pong , and a Slot Machine Simulator using Pygame, Tkinter, NumPy , and Turtle graphics . Implemented features like GUI rendering, event-driven programming, collision detection, control flow logic , and real-time input processing to simulate interactive gameplay and adaptive interactions.	Sep 2024 - Dec 2024
--	---------------------