Hasini Rasagna Shakamuri

Hyderabad, India | +91 9494196777 | hasinirasagna@gmail.com | LinkedIn | GitHub

EDUCATION

Master of Engineering in Computer Science	**Aug 2022 – Dec 2023**
University of Cincinnati, Cincinnati, Ohio	GPA: **3.9/4**
Bachelor of Technology in Computer Science	**July 2014 - April 2018**
MRCET, Hyderabad, India	GPA: **8.2/10**

TECHNICAL SKILLS

Cloud Based Technologies	Snowflake, AWS, Microsoft Azure, IICS, GCP, Databricks
Computing DW Platforms	Snowflake, redshift, Azure Synapse, Teradata, IBM DB2
ETL/ELT Tools	Informatica power centre, IBM Data Stage, Enterprise Service Platform, AbInitio
BI Reporting Tools	Looker, Tableau, ESP Reports, Power BI
Languages/ Scripting	Python, PySpark, PL/SQL, UNIX, Shell Scripting
Databases	IBM DB2, MySQL, Oracle, SQL Server, Hive, PostgreSQL, Sybase, MongoDB, DynamoDB, Teradata, NoSQL, SSIS
Tools & Utilities	Erwin, SAP Power Designer, Jenkins, GitHub, Autosys, DBT, Control M, JIRA, Apache Airflow, NiFi, Swagger portal, Docker, Control -M, CI/CD pipeline

PROFESSIONAL EXPERIENCE

Data Analytical Engineer	**June 2024 – Nov 2024**
Erie Insurance	**PA, United States**

Environment: Snowflake, IBM DB2, SQL, Python, Data Modelling (Logical & Physical), Code/Decode, Data Mapping, Data Integration, Data Migration, Enterprise Data Standards, UNIX, JIRA

- Translated business and third-party insurance data requirements into logical and physical data models, aligning DB2 source systems with Snowflake target schemas and enforcing enterprise data standards.
- Performed code/decode transformations and canonical mapping on DB2 datasets to normalize external insurer attributes before integrating them into downstream analytical databases.
- Designed and executed SQL-based data migration and validation scripts to move and reconcile data across DB2, Snowflake, and relational platforms, ensuring consistency and historical accuracy.
- Partnered with third-party data providers to analyze source schemas, understand data semantics, and define column-level mappings, supporting seamless integration into Snowflake pipelines.
- Supported data ingestion and integration pipelines by assisting migration and development teams with schema clarifications, data lineage, and validation logic, acting as the data SME during onboarding.

Cloud Engineer	**Jan 2023 – June 2024**
Top Stars Inc	**Atlanta GA, USA**

Environment: Snowflake (Stages, Snow Pipes, Snowpark, External Tables, Time Travel, Fail-safe, SnowSQL, Task Scheduling), Python, PySpark, Informatica, SQL Server, Looker, PowerBI, Git, Apache Airflow, Axiom, Tableau, AWS S3, SNS, SQS, CloudWatch, Batch & NRT data streaming, structured and Unstructured data

- Create efficient and scalable data models and pipelines within Snowflake, considering performance optimization and best practices.
- Own and manage the data quality and data integrity of the 2nd most queried table in the data pipeline.
- Develop Snowflake external **stages** and build **Snow Pipes** to enable **near real-time** data processing from **AWS S3** to Snowflake staging tables, optimizing data ingestion workflows.
- Implement Snowflake **Time Travel** and **Fail-safe** mechanisms to ensure data integrity and facilitate seamless data recovery.
- Establish secure **data sharing protocols** within Snowflake environment, enhancing collaboration and data accessibility.
- Configure **AWS SQS** service for efficient file triggering from S3, streamlining data loading processes into Snowflake.
- Specialized in performance optimization and tuning techniques within Snowflake via various SQL queries.
- Design and implement **Snowflake data warehouses and data lakes**, aligning architecture with business requirements.

Consultant   **August 2022 – Dec 2023**
University of Cincinnati   **OH, United States**

- Independently supervised class of 30-40 and cleared doubt in SQL concepts and explain the way to optimize SQL queries along with assisting them in SQL Lab.
- Conducted quizzes focusing on emphasizing key methods resulting in a noticeable improvement in students' performance.
- Managed software installations, conducted rigorous testing procedures, and resolved technical issues for students in the CEAS department at UC.

Data Engineer   **June 2021 – July 2022**
JP Morgan Chase & Co   **Hyderabad, India**

Environment: Apache Spark, PySpark, Python, SQL, Data warehouse, RDD Transformations, AWS, Redshift, Athena, Glue, AWS lambda, RDS, EMR, ETL, Bitbucket, controlM, JIRA, Pipeline ingestion, UNIX, OLTP, OLAP, HDFS, Macros, Excel

- Handled multiple LOBs, **ingested data** from various data sources to HDFS. Wrote **UNIX** scripts to automate the registration and ingestion process, streamlining data import into the target system. Maintained design document throughout the project execution.
- Designed & implemented datasets**,** data **pipelines** (**JSON**) and registered in swagger portal, generating a GUID.
- Developed an **ETL** logic with a predefine **transformation** logic and triggered them with **API** to load data into a parquet file.
- Performed PRANA testing by designing over 350+ test cases in **PySpark** to validate the source file with the target HDFS file.
- Executed **controlM** jobs & workflows for performance, scalability & automated the ingestion process in higher environments.
- Enforced code version control using **bitbucket** & promoted to UAT, BUAT & PROD. Supervised defect tracking and tasks on Atlassian **JIRA**. Successfully worked on production release and deployed 300+ tables to desired databases.
- Successfully migrated data from **HDFS** to **Amazon S3, Snowflake Virtual warehouse** and performed ETL using **AWS Glue,** streamlining data import into the data lake. Implemented **AWS Lambda** for event-driven automation of data ingestion.
- Designed data pipelines and implemented **ETL logic** using **AWS Glue** and triggered these jobs utilizing **AWS Lambda** to load data into Parquet files. Performed thorough data validation by querying data using **AWS Athena**, guaranteeing data integrity.

Data Engineer,   **June 2018 – June 2021**
State Street Corporation   **Chennai, India**

Environment: Azure Databricks, Data Lake, Blob Storage, SQL DB, Azure Data Factory, Cluster Pool, Key Vault, Agile, Python

- Utilized Python, PySpark and SQL to construct automated data pipelines capable of handling the humongous data. The pipelines transformed transactional data into fact and cube tables to support analysis and facilitate low latency dashboards.
- Led work to design and implement release best practices for data pipelines (resources, permissions, monitoring, storage, testing, scheduling, and compliance) which lowered production outages by 30% and doubled weekly releases.
- Applied the spark DataFrame API within spark session to implement advanced data manipulations and analysis efficiently.
- Created an executive-level one-pager dashboard, effectively presenting operational metrics to top-level management.

Data Analyst

Environment: PL/SQL, Oracle, SQL, Git, Tableau, Data visualization, ETL, ESP Enterprise Solution Platform, Reporting, T-SQL, Excel, Macros, Agile, OLAP, Excel, UNIX

- Researched & analysed portfolio data to build over **100 tables, views, schemas, cursors** as applicable to transformation and generate reports to support **data driven decision making**.
- Extracted data from **ESP ETL** and created **80+** reports and added new functionalities for the existing reports as clients specified by writing **window functions, joins, subqueries**. Triggered and scheduled **100+** tables to pull the data and generate reports every day.
- Developed & maintained **PL/SQL packages, procedures, functions, triggers, and Indexes** to manage data and generate reports.
- Conducted **root cause analysis** and implemented effective **problem-solving** techniques to address data-related issues, used

performance tuning and improved performance and scalability. Pushed code to **GIT** and managed version control.
- Analysed data applying statistical techniques and rendered effective **visualizations** in form of **reports, charts, graphs.**
- Executed iterative report generation utilizing Job scheduling for both legacy & transactional data & **reduced runtime** by 25% and Enhanced Reports for scalability, Enhanced data delivery which has **improved performance** by 10%.

Business Intelligent Analyst, IBM   **India, May 2017 – April 2018**
- Studied data warehouse requirements and worked on all phases from gathering requirements to testing, implementation, data migration and support. Presented PowerPoint briefings and executive summaries to line of business executives.

Certifications

AWS Certified Cloud Practitioner - Certification Number: EWLDKBCLYF44QTGF
Microsoft Certified Azure Fundamentals - Certification Number: I704-4849
Microsoft Certified Azure Data Engineer Associate - Certification Number: 2J4465-2CCB5A

ACADEMIC PROJECTS

Projects: GithubLink
File Uploader | Amazon EC2, S3, IAM, DynamoDB, Python, HTML, CSS, Flask
- Built intuitive and user-friendly interfaces facilitating secure storage of information in cloud for seamless file uploads.

Enquire Bot | Python, Google Cloud Dialog Flow, Natural Language Processing
- Designed and programmed a college enquiry chatbot applying various Python-based tools, modules, and Google Cloud Dialogflow.

IP&TEXT PARSER | Docker, Python
- Analysed text data and extracted IP addresses through Python and its associated libraries, with containerization handled via Docker.

Retail Data Application | Azure Cloud
- Created an interactive web application in Azure cloud deploying real-world retail data from 84.51°/Kroger and applying the azure technology stack for innovative problem-solving solutions to improve customer engagement.

Big Data | PySpark using Anaconda & Jupyter notebook, Databricks, Data Visualization
- Conducted exploratory data analysis adapting Databricks Notebook on crime data of 3 major cities with PySpark, SQL, and Data Visualization techniques to gain insights into crime patterns and trends.

LEADERSHIP EXPERIENCE AND RECOGNITIONS

- Won District and State Level Chess Competitions.
- Founder of IN-Need NGO, an initiative to Nurture, Educate, Encourage and Donate unprivileged people.
- Received Sliver Certification from Industrial AGILE Office for following Industrial AGILE Methodologies.
- Achieved many SPOT Recognition awards for excellence in performance during Work Experience.