SANDEEP PANDELLAPALLI

📞 217-904-2989 ✉ sandeep.pandellapalli@gmail.com in linkedin.com/in/prabhakarasandeep GitHub 📍 Urbana, IL - 61801

Education

University of Illinois Urbana Champaign   **Aug 2023 – May 2025**
Master of Science in Information Management   *Champaign, Illinois*
- Relevant Coursework: Information Modeling, Data Warehousing & Business Intelligence

Gitam University   **Jun 2017 – May 2021**
Bachelor of Science in Mechanical Engineering   *Visakhapatnam, India*
- Relevant Coursework: Probability & Statistics, Database Management

Skills

- **Programming & Scripting Languages:** Python, SQL, Unix, PySpark
- **Data Warehousing & ETL Tools:** Oracle, PostgreSQL, Git, Informatica, Snowflake
- **Cloud & Big Data Technologies:** AWS (S3, Redshift, Glue, Athena, Lambda), Apache(Spark, Kafka, Airflow)
- **BI Tools:** PowerBI, Tableau, AWS Quicksight
- **Certifications:** Amazon - **AWS Data Engineering Associate**(DEA-C01), Snowflake - **SnowPro Core(COF-C02)**

Experience

Gies College of Business   **Nov 2023 – Present**
Research Assistant   *Champaign, Illinois*
- **Bridged the gap** between researchers and business clients by translating technical requirements into data solutions aligned with business goals.
- Directed a team to create a **serverless AWS-based RAG chatbot** for a Fortune 500 company, optimizing the architecture to cut search times from two minutes to under five seconds.
- Leveraged **Snowflake's Snowpipe and Streams** to refine historical data analysis, increasing the depth of trend insights by 30% and enhancing the prediction of Customer Lifetime Value (CLV).

Cognizant   **Jul 2021 – Jul 2023**
Data Engineer   *Bangalore, Karnataka*
- Pioneered **12 ETL projects** in the Analytics department orchestrating requirements gathering, data loading strategies, and development of mappings and workflows.
- Spearheaded the client's **outdated business systems transformation** to build an Event-Driven Architecture (Kafka/Flink/Airflow) saving $300K in annual operational efficiencies.
- Partnered with QA and product teams to implement an **automated data validation framework** using PL/SQL and Python, resolving 200+ schema mismatches.
- Formulated ETL pipelines **using AWS Glue and PySpark** enforcing seamless processing of datasets and implementing optimizations that slashed query execution time by 50%.

Cognizant   **Feb 2021 – Jun 2021**
Data Engineer Intern   *Bangalore, Karnataka*
- Engineered and deployed **6 Linux shell scripts** to streamline Informatica workflows, improving data transformation while leveraging cloud infrastructure to ensure scalability and efficiency in data processing.
- Migrated **20 complex datasets** into a unified Oracle SQL database, improving data accessibility for cross-functional teams and eliminating three critical data silos that slowed decision-making.
- Spearheaded the deployment of **Slowly Changing Dimension Types 1 and 2** for the Clinical Trials - Risk Assessment Categorization Tool (CT-RACT) project, enhancing data accuracy by 30% .
- Enhanced **data-driven decision making** by engineering scalable No-SQL data pipelines using Amazon DynamoDB, reducing data retrieval times by 35%.

Academic Projects

Automated Big Data Pipeline: Spark Bulk Data Load with CI/CD Integration   **Jan 2025**
- Engineered a high-performance Spark-based data pipeline with a **Jenkins-driven version control** workflow to extract, transform and load entity records from Hive to Kafka, reducing data processing time by 50%.

Financial Trend Analysis: McDonald's Post E. coli Outbreak   **Nov 2023**
- Analyzed McDonald's financial performance post E. coli outbreak by transforming SEC datasets with Python and developing a Tableau dashboard, enhancing predictive accuracy of financial trends by 35%.

Real-Time Data Integration: Change Data Capture across Cloud Services   **Nov 2024**
- Deployed a Kafka based event driven data transport layer using Confluent Cloud, streamlining **MySQL database changes to AWS DynamoDB and AWS Lambda.**