

Anusha Baradi
Data Engineer
+1 (848) 306-5100 | anushabaradi25@gmail.com

Available for C2C roles only

PROFESSIONAL SUMMARY:

8+ years of strong experience in data modeling, processing, analysis and extensively worked on processes to resolve complex business problems involving large scale structured, semi-structured data.

Technical skills:

RDBMS	:	Snowflake, Redshift, PostgreSQL, SQL Server
ETL	:	Pyspark, SSIS
Big Data	:	Spark, EMR, Data Lake, Elasticsearch
Scripting	:	Python
Cloud Technologies	:	AWS
Repository Tools	:	GIT

Education:

Bachelor of Science in Computer Science & Engineering 2011.

Professional Experience:

USAA (April 2022 – Till Date)

Role: Data Engineer

Location: San Antonio, Texas (Remote)

Responsibilities:

- Extensively worked on data ingestion, pipeline design, data modeling, data mining and advanced data processing.
- Designed and implemented data pipelines to process semi-structured and structured data using different python libraries like pyspark, boto3, pandas and SQL Alchemy.
- Experienced in extracting batch data from different data sources like Parquet/Text/CSV Files from S3, databases, and applied transformations as per business rules and stored data into Snowflake, Redshift and PostgreSQL.
- Experienced in working on cloud-based data solutions using different AWS Services like EMR, S3, DMS, Data Lake, Lambda, Event Bridge, CloudWatch, Glue, DynamoDB, Redshift, IAM, EC2, RDS and Athena.
- Developed python scripts to automate data sampling process and ensured data integrity by checking for completeness, duplication, accuracy, and consistency.
- Built and optimized complex SQL queries involving multiple joins to perform data extraction and merging from large volumes of historical data.
- Compared data between various platforms and debugged the code to find the breaks in data and fixed them.
- Designed unit test cases to ensure the code performs as expected.
- Involved in understanding existing business requirements based out of star, snowflake data models and logical, physical data.

- Extensively worked in Phantom / Splunk SOAR, including writing playbook, troubleshooting, training, or supporting technical requests.
- Involved in writing queries in Elasticsearch and monitored the logs.
- Automated tasks of a security analyst using python and Splunk Phantom by gathering the requirements and implementing them.
- Assisted scrum master in sprint planning sessions by providing backlog items and story points, reviewed team user stories and acceptance criteria.
- Managed project versions and source code changes using source control tools GIT.

ProSight Specialty Insurance (Jan 2020 – Mar 2022)

Role: Python Developer

Location: Morristown, NJ

Responsibilities:

- Involved in analysis, specification, design, and implementation and testing phases of Software Development Life Cycle (SDLC) and used agile methodology for developing data delivery applications.
- Built database models in PostgreSQL on AWS, wrote queries and stored procedures for various data retrievals.
- Designed python applications using psycopg2 to interface with PostgreSQL database and perform CRUD operations.
- Developed various PostgreSQL database queries from python using Python- PostgreSQL connector and PostgreSQL DB package.
- Created python scripts to handle Captured images on filesystem from DB backup.
- Developed data priming scripts, data migration scripts, metadata backups to csv files using python csv.
- Designed PySpark Data frames to handle data from the client facing source systems, routed the data to the DB tables.
- Create JSON files using Python JSON, to deliver data to downstream system's dashboard applications.
- Developed Python scripts to handle data changes in database tables for data fixes and bulk updates.
- Designed and developed python applications to handle data from files and generate PDF files.
- Designed and developed Python OOD code for quality, logging, monitoring, and debugging code optimization.
- Development of Python APIs to dump the array structures in the Processor at the failure point for debugging.
- Heavily involved with team, collaborated in an Agile Development workflow setup and used Git for version control.
- Developed test cases and test scripts using unit test to ensure that code performs as designed.
- Worked with testing and production support teams to help in code deployment, also developed backout scripts, handled code backups. Used JIRA for Bug tracking and issue tracking.
- Worked as part of an Agile/Scrum based development team.

Premier Healthcare (Jul 2018 – Nov 2019)

Role: Python Developer

Location: Charlotte, NC

Responsibilities:

- Designed and developed tables to capture new data sets. Developed python classes to handle data specific to region and client. Data from the UI is captured into their tables, with backend classes.
- Performed the daily runs for claims engine for various customers, deployed automatic runs as well as manual runs to facilitate the process. Also Created process to audit and archive data backups.
- Extensively wrote SQL to interact with SQL Server database, for the code data sourcing the claims data engine.
- Developed data processing applications in Python to deliver consolidated data to the UI and Reporting systems.
- Extensively used Psycpg2, IO, Pandas libraries to gather data from the database and apply the business rules to further the data processing.
- Developed applications to supply data for the Pyspark engine for processing, worked with source data in Hive database, with file format such as Paraquet.
- Developed testing scripts unit test, designed and stored test cases as part of test-driven development.
- Worked with code deployment teams, release management and product support teams to help in Bug tracking and issue tracking. And post deployment application support.
- Logged user stories and acceptance criteria in JIRA for features by evaluating output requirements and formats.
- Fixed bugs, providing production support, enhanced applications by improving code reuse and performance by making effective use of various design patterns.

AT&T (May 2016 – June 2018)**Role: Backend Developer****Location: Bedminster, NJ****Responsibilities:**

- Involved in analysis, specification, design, and implementation and testing phases of Software Development Life Cycle (SDLC) and used agile methodology for developing application.
- Participate in requirement gathering and analysis phase of the project in documenting the business requirements by conducting workshops/meetings with various business users.
- Worked on optimizing and memory management of the ETL services.
- Performed Data mapping between source systems to Target systems, logical data modeling, created class diagrams and ER diagrams and used SQL queries to filter data.
- Wrote and executed various MYSQL database queries from Python using Python-MYSQL connector and MySQL db package.
- Used PySpark-SQL to load JSON data and create schema RDD, Data Frames and loaded it into Hive Tables and handled Structured data using Spark-SQL.
- Utilized Python Libraries like Boto3 for AWS. Used Pandas library for statistical Analysis.
- Worked with team of developers on Python applications for RISK management.
- Using Git version control tool to coordinate team-development.
- Implemented PySpark using Python with data frames and temporary table SQL for faster data processing.
- Worked on several python packages like Pandas, SQL Alchemy, and sockets.
- Responsible for debugging the project monitored on JIRA (Agile).

Lamda (May 2011 to Apr 2014)**Role: Python Developer**

Location: India**Responsibilities:**

- Responsible for analyzing various applications systems enforcing python, best practices and provide guidance in making long term architectural design decisions.
- Involved in entire lifecycle of the projects including Design, Development, and Deployment, Testing and Implementation and support.
- Debugging Software for Bugs.
- Generated property list for every application dynamically using Python.
- Developed ETL scripts based using Python framework.
- Implemented Test Driven Development (TDD) strategy for the project.
- Developed Python batch processors to consume and produce various feeds.
- Responsible for debugging the project monitored on JIRA (Agile).
- Built development environment with JIRA/Git.