Gangishetty S

Data Engineer, 10+ years experience
Hyderabad, India
https://augmntx.com/profile/data-engineer-gangishetty-s/axf29112

AugmntX
+91 9820045154 | hr@augmntx.com

Data Engineer with 9.6 years of experience in designing and implementing scalable data solutions that drive informed, data-driven decision-making. Skilled in building robust ETL pipelines, data warehousing architectures, and end-to-end data processing workflows. Strong hands-on expertise in SQL, Python, and PySpark for transforming large datasets and optimizing pipeline performance. Adept at solving complex business problems through efficient data modeling and automation. Experienced in working with cross-functional teams to deliver reliable, high-quality data solutions. Known for analytical thinking, attention to detail, and a strong commitment to scalability and performance.

INDUSTRIES: Finance, Insurance, Manufacturing

TECHNICAL SKILLS

• SQL	9 years & 10 months
• Azure Data Bricks	5 years
• Python	5 years
• Pyspark	5 years
• Azure Data Factory V2	4 years
• Azure Synapse Analytics	3 years

PROJECT DETAILS

• Market Data Hub — April-2023 - June-2025

Project Description: Market Data Hub is a Data Engineering team that on boards different vendors raw data, cleanse the data and load it to azure data lake. Business users will consume the data for business decisions in making of investments.
Responsibilities: Collaborated with data vendors and business users to understand their specific needs. Designed and orchestrated end-to-end ETL for data movement and transformation workflows using Azure Synapse pipelines, Azure Databricks (Pyspark) and integrating CI/CD for streamlined production deployment. Developed PySpark codes for testing to perform data quality checks and validate data.
Technologies: Azure Databricks(Pyspark)), Python, Azure Synapse Analytics, Azure Sql Database,SQL,ADLS
Industry: Finance

• FFWR August-2015 - June-2019

Project Description: A automate the process of Life and Annuity of Premium for users to know their individual policy life, agents incentives and commission through their web reports. It is a daily scheduled process to know about agent and individual customer and agencies information. This boosts the Insurance domain in various ways.

Responsibilities: Collaborated with cross-functional teams and business stakeholders to understand the requirements and align technical solutions with business requirements. Developing ETL Queries to generate data in specified format for loading through SSIS and Data Bricks Extensively used ETL Components to Extract Data from Files, Coding & Integration of the application to other applications. Performing transformation over the extracted data to cleanse the file. Cleansing was done using spark SQL Queries. Cleaning operations such as removing null, bad data etc. were performed Additions of new columns such as audit columns using Pyspark were implemented. Created Pipelines in Data Factory related to Databricks Notebook for complete end-to-end flow of functionality. Responsible in performing sort, join, aggregations, filter, and other transformations on the data frames using Spark. Handle the client calls to procure the business requirements of the support, enhancement and development related work of the project.

Technologies: Azure Databricks(Pyspark)), Python, Azure Data Factory, Azure Sql Database,SQL,ADLS,SSIS,Azure Synapse Analytics

Industry: Insurance

• Rapid Radical July-2021 - April-2023

Project Description: Rapid Radical is Data Engineering team that receives TBs of raw data of company products, cleanse the data and load it to azure data lake. Data scientists will leverage this platform to discover data, conceptualize solution, build & test models, and deploy in production. Business users will consume the data for business decision making to achieve cost reduction and achieve operational efficiency

Responsibilities: Collaborated with cross-functional teams and business stakeholders to understand the requirements and align technical solutions with business requirements. Designed and orchestrated end-to-end ETL for data movement and transformation workflows using Azure Data Factory pipelines, Azure Databricks (PySpark). Implemented Spark using Python in data bricks utilizing Data Frames and Spark-SQL API for faster Processing of Terabytes of data. Developed PySpark codes for testing to perform data quality checks and validate data.

Technologies: Azure Databricks(Pyspark)), Python, Azure Data Factory, Azure Sql Database,SQL,ADLS

Industry: Manufacturing

EMPLOYMENT HISTORY

• Senior Software Developer April-2023 - June-2025

Description: • Collaborated with data vendors and business users to understand their specific needs. • Designed and orchestrated end-to-end ETL for data movement and transformation workflows using Azure Synapse pipelines, Azure Databricks (Pyspark) and integrating CI/CD for streamlined production deployment. • Developed PySpark codes for testing to perform data quality checks and validate data.

• Technical Lead July-2021 - April-2023

Birlasoft

Description: Technical Lead

CERTIFICATION

• Databricks Certified Data Engineer Associate by Databricks 2025

• Microsoft Certified Azure Data Engineer Associate by Microsoft 2020

EDUCATION

• Bachelors degree in ECE from Osmania University 2011 - 2015

LANGUAGES

• English - advanced

END OF DOCUMENT
This PDF is automatically generated by AugmntX on 21 Nov 2026
Profile URL:https://augmntx.com/profile/data-engineer-gangishetty-s/axf29112