

Sagnik Adusumilli

Big Data engineer with over two years of experience in ETL and Ingestion

#15-30 Chichester Pl

Toronto, M1T 3S5

(647) 978-2744

sagnik.adusumilli@gmail.com

EXPERIENCE

TD Bank, Toronto— *Lead Data Engineer*

July 2022 - PRESENT

Lead Data engineer since September 2024

Lead a team of offshore/onshore engineers for developing a tool using PySpark and Azure Data Factory (ADF) to automate rollback procedures for production deployment failures. Rollback process involved removing data on cloud, reverting delta table schema and reverting SSMS table changes.

Designed and developed an end-to-end automation tool using PySpark and Databricks workflows to monitor ingestion jobs and trigger DEV unit testing. This saved a significant amount of manual effort and removed human errors.

ETL - Data cleaning, transformations and automations done using PySpark and ADF pipelines
Provided business insights on data such as data formats, load types (incremental, transactional etc.) in, data security and investigated data processing issues.

ETL - Extracted and ingested data in formats such as delimited, Json, API, database, etc., to Standardized Raw Zone (SRZ) as delta tables stored in Parquet file format.

ETL (On prem) - Developed stored procedures in SQL to extract and load data in star schema format and implemented Slow Changing Dimensions (SCD) type 2.

Used ADF pipelines to run various jobs such as ingestion tasks, set up schedulers, modify SSMS tables using JDBC drivers etc.

Resolved 20+ (highest in the team) critical production issues - some with record turnaround times. Helped team members in resolving their issues.

Aviva, Toronto— *Software Developer Intern*

May 2018 - August 2019

Worked as a full-stack developer on an online portal used by underwriters/brokers to create insurance quotes for small businesses. (Tech stack: NodeJS, Angular, MongoDB, Handlebars.js)

Presented Data science and machine learning methods during showcase event

SKILLS

Azure Data Factory, Apache spark (PySpark and Spark SQL) on Databricks for Big Data ingestion and ETL

SQL MySQL DBMS, PostgreSQL DBMS and SQL Server Management Studio (SSMS)

Delta Tables

Python, Java, C, SQL

Git, Jira and Bitbucket

PROJECTS (More on [Github](#))

F1 ETL

End to End data engineering project that implements medallion architecture and orchestration to ingest and present F1 data to a dashboard

Bike Toronto ETL

ETL and data analysis project done using PostGreSQL to draw meaningful insights for cyclists in Toronto

Covid Data analysis

A study evaluating the effectiveness of the COVID-19 vaccine across various strains and age groups

EDUCATION

University of Toronto, St George — B.Sc

November 2022

Math, Statistics and Computer Science

AWARDS/Certifications

Databricks Lakehouse Fundamentals

Microsoft Certified: Azure Fundamentals

Multiple time “Above and beyond” award recipient at TD (2024)

Aviva Hackathon Runner up (2019)

LinkedIn: <https://www.linkedin.com/in/sagnik-adusumilli-645352124/>

GitHub: <https://github.com/SagnikAdusumilli>